

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 18:29:19 ; Search time 224 Seconds
(without alignments)
3205.659 Million cell updates/sec

Title: US-10-646-396-2

Perfect score: 6619
Sequence: 1 MDKARNCLLQHREALEKDI.....PKTYVTVDNLGILYLQTL 1248

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6619	100.0	1248	1 APAF HUMAN	O14727 homo sapien
2	5873.5	88.7	1249	1 APAF MOUSE	O88879 mus musculus
3	5871.5	88.7	1249	1 APAF RAT	Q98p55 rattus norv
4	5863.5	88.6	1249	2 Q8VI66	Q8vi66 rattus norv
5	4054	61.2	1248	2 Q8GNJ6	Q8gnj6 xenopus lae
6	3683.5	55.7	1261	1 APAF BRARE	Q919h8 brachydanio
7	1066	16.1	258	2 Q80VR5	Q80vr5 mus musculus
8	732	11.1	1227	1 Q8ZOR1	Q8zor1 anabaena sp
9	728.5	11.0	1258	1 Y800 ANASP	Q8ytc2 anabaena sp
10	693	10.5	1184	2 Q7ND85	Q7nd85 gloeobacter
11	691	10.4	1526	1 YK46 ANASP	Q8yri1 anabaena sp
12	689.5	10.4	1197	2 Q7NJ67	Q7nj67 gloeobacter
13	677.5	10.2	1711	2 Q8Z019	Q8z019 anabaena sp
14	672	10.2	1683	1 Y124 ANASP	Q8yv57 anabaena sp
15	666	10.1	1193	2 Q7ND05	Q7nd05 gloeobacter
16	659	10.0	1188	2 Q7ND80	Q7nd80 gloeobacter
17	643.5	9.7	1356	1 HET1 PODAN	Q00808 podospora a
18	626	9.5	1693	1 Y163 SYN3	Q55563 synchocyst
19	624.5	9.4	1356	2 Q8X1P4	Q8x1p4 podospora a
20	614.5	9.3	934	2 Q8Y223	Q8yz23 anabaena sp
21	611.5	9.2	1551	2 Q8YMU3	Q8ymu3 anabaena sp
22	610	9.2	1708	2 Q8Y212	Q8yzi2 anabaena sp
23	609	9.2	1081	2 Q7NCT8	Q7nct8 gloeobacter
24	606.5	9.2	1356	2 Q8X1P3	Q8x1p3 podospora a
25	605.5	9.1	1356	2 Q8X1P2	Q8x1p2 podospora a
26	604.5	9.1	1376	2 Q8Y109	Q8y109 anabaena sp
27	600.5	9.1	1189	2 Q8Y109	Q8y109 anabaena sp
28	586	8.9	1183	2 Q7NLE9	Q7nle9 gloeobacter
29	581	8.8	1233	2 Q8TMX4	Q8tmx4 methanosarc
30	580.5	8.8	1189	2 Q8YTD1	Q8ytd1 anabaena sp
31	561.5	8.5	1747	2 Q8Z020	Q8z020 anabaena sp

32	539	8.1	1241	2 Q9XBD8	Q9xbd8 amycolatops
33	531.5	8.0	1730	2 Q7NKS0	Q7nk50 gloeobacter
34	530.5	8.0	1193	2 Q7NH82	Q7nh82 gloeobacter
35	508	7.7	1051	2 Q8TMS3	Q8tms3 methanosarc
36	506	7.6	1671	2 Q7NMP0	Q7nmp0 gloeobacter
37	493.5	7.5	1721	2 Q7NID9	Q7nid9 gloeobacter
38	464.5	7.0	1329	2 Q8YYT7	Q8yyt7 anabaena sp
39	444	6.7	1538	2 Q6W219	Q6w219 rhizobium s
40	444	6.7	1538	2 AAQ87199	AAq87199 rhizobium
41	442.5	6.7	1592	2 O62471	O62471 caenorhabdi
42	442.5	6.7	1592	2 CAAL6357	Caal6357 caenorhab
43	438	6.6	1237	2 Q6ZE54	Q6ze54 synchocyst
44	438	6.6	1237	2 BAD02046	Bad02046 synchocyst
45	423	6.4	2629	1 TEPI_RAT	O08653 rattus norv

ALIGNMENTS

RESULT 1

ID	APAF HUMAN	STANDARD	PRT	1248 AA
AC	O14727; Q7Z438; Q9BXZ6; Q9UBZ5; Q9UGN8; Q9UGN9; Q9UGP0; Q9UJ58; Q9UJ59; Q9UJ60; Q9UJ61; Q9UJ62; Q9UJ63; Q9UJ64; Q9UJ65; Q9UJ66; Q9UJ67; Q9UNC9;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Apoptotic protease activating factor 1 (Apaf-1).			
GN	Name=APAF1; Synonyms=KIAA0413;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 3), AND PARTIAL SEQUENCE.			
RC	TISSUE=Cervical carcinoma;			
RX	MEDLINE=97410306; PubMed=9267021;			
RA	Zou H., Henzel W.J., Liu X., Iutischg A., Wang X.;			
RT	"Apaf-1, a human protein homologous to C. elegans CED-4, participates in cytochrome c-dependent activation of caspase-3.";			
RL	Cell 90:405-413(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1; 4 AND 5).			
RC	TISSUE=Cervical carcinoma, Heart, and Peripheral blood;			
RX	MEDLINE=99373149; PubMed=10441496; DOI=10.1006/bbrc.1999.1124;			
RA	Hahn C., Hirsch B., Janke D., Duerkop H., Stein H.;			
RT	"Three new types of Apaf-1 in mammalian cells.";			
RL	Biochem. Biophys. Res. Commun. 261:746-749(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=T-cell;			
RX	MEDLINE=99292765; PubMed=10364241;			
RA	Saleh A., Srinivasula S.M., Acharya S., Fishel R., Alnemri E.S.;			
RT	"Cytochrome c and dATP-mediated oligomerization of Apaf-1 is a prerequisite for procaspase-9 activation.";			
RL	J. Biol. Chem. 274:17941-17945(1999).			
RN	[4]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, SUBUNITS, AND MUTAGENESIS OF			
RC	LYS-160 AND MET-368.			
RX	TISSUE=Kidney;			
RA	Hu Y., Benedict M.A., Ding L., Nunez G.;			
RT	"Role of cytochrome c and dATP/ATP hydrolysis in Apaf-1-mediated caspase-9 activation and apoptosis.";			
RL	EMBO J. 18:3586-3595(1999).			
RN	[5]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	Ohara O.;			
RA	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE OF 385-1248 FROM N.A. (ISOFORM 2).			
RC	TISSUE=Brain;			

RX MEDLINE=98116655; PubMed=9455477;
RA Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VIII.
RT 78 new cDNA clones from brain which code for large proteins in
RT vitro.";
RL DNA Res. 4:307-313(1997).
RN [7]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 6), FUNCTION, AND SUBCELLULAR
RP LOCATION.
RC TTSSUE=Prostatic carcinoma;
RX MEDLINE=2268499; PubMed=12804598; DOI=10.1016/S0006-291X(03)00995-1;
RA Ogawa T., Shiga K., Hashimoto S., Kobayashi T., Horii A., Furukawa T.;
RT "APAF-1-ALT, a novel alternative splicing form of APAF-1, potentially
RT causes impaired ability of undergoing DNA damage-induced apoptosis in
RT the LNCaP human prostate cancer cell line.";
RL Biochem. Biophys. Res. Commun. 306:537-543(2003).
RN [8]
RP SEQUENCE OF 810-864 AND 866-883 FROM N.A.
RA Roberts D.L., Dalgleish R., Cohen G.M., MacFarlane M.;
RT "the mammalian CED4 homologue, APAF1, exists as two distinct forms in
RT human cells.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE OF 1-138 FROM N.A. (ISOFORM 1/4/5).
RA Won M., Lee J.-W., Oh H.-H., Kim D.-U., Chung K.-S., Lee M.,
RA Yoo H.-S.;
RT "Cloning of variant Apaf1.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [10]
RP APAF-1-MEDIATED OLIGOMERIZATION.
RX MEDLINE=98315378; PubMed=9651578;
RA Sriniwasula S.M., Ahmad M., Fernandes-Alnemri T., Alnemri E.S.;
RT "Autoactivation of procaspase-9 by Apaf-1-mediated oligomerization.";
RL Mol. Cell Biol. 19:949-957(1999).
RN [11]
RP INDUCTION BY E2F AND P53.
RX MEDLINE=21283226; PubMed=11389439; DOI=10.1038/35078527;
RA Moroni M.C., Hickman E.S., Denchi E.L., Caprara G., Colli E.,
RA Cecconi F., Mueller H., Helin K.;
RT "Apaf-1 is a transcriptional target for E2F and p53.";
RL Nat. Cell Biol. 3:552-558(2001).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS) OF 1-97.
RX MEDLINE=20013059; PubMed=10543941; DOI=10.1006/jmbi.1999.3177;
RA Vaughn D.E., Rodriguez J., Lazebnik Y., Joshua-Tor L.;
RT "Crystal structure of Apaf-1 caspase recruitment domain: an alpha-
RT helical Greek key fold for apoptotic signaling.";
RL J. Mol. Biol. 293:439-447(1999).
RN [13]
RP STRUCTURE BY NMR OF 1-97.
RX MEDLINE=20047184; PubMed=10578182;
RA Day C.L., Dupont C., Lackmann M., Vaux D.L., Hinds M.G.;
RT "Solution structure and mutagenesis of the caspase recruitment domain
RT (CARD) from Apaf-1.";
RL Cell Death Differ. 6:1125-1132(1999).
RN [14]
RP FUNCTION: Oligomeric Apaf-1 mediates the cytochrome c-dependent
RP autocatalytic activation of pro-caspase-9 (Apaf-3), leading to the
RP activation of caspase-3 and apoptosis. This activation requires
RP ATP. Isoform 6 is less effective in inducing apoptosis.
CC [15] SUBUNIT: Monomer. Oligomerizes upon binding of cytochrome c and
CC dATP. Oligomeric Apaf-1 and pro-caspase-9 bind to each other via
CC their respective N12-terminal CARD domains and consecutively
CC mature caspase-9 is released from the complex. Pro-caspase-3 is
CC recruited into the Apaf-1-pro-caspase-9 complex via interaction
CC with pro-caspase-9.
CC [16] SUBCELLULAR LOCATION: Cytoplasmic.
CC [17] ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=6;
CC Name=1 Synonyms=Apaf-1X1;
CC IsoId=O14727-3, Sequence=displayed;
CC Name=2, Synonyms=Apaf-1L, Sequence=displayed;
CC IsoId=O14727-2, Sequence=VSP_006759;
CC
CC Name=3; Synonyms=Apaf-1S;
CC IsoId=O14727-3; Sequence=VSP_006759, VSP_006761;
CC Name=4; Synonyms=Apaf-1M;
CC IsoId=O14727-4; Sequence=VSP_006761;
CC Name=5; Synonyms=Apaf-1XS;
CC IsoId=O14727-5; Sequence=VSP_006760, VSP_006761, VSP_006762;
CC Name=6; Synonyms=Apaf-1-ALT;
CC IsoId=O14727-6; Sequence=VSP_008965, VSP_008966;
CC [18] TISSUE SPECIFICITY: Ubiquitous. Highest levels of expression in
CC adult spleen and peripheral blood leukocytes, and in fetal brain,
CC kidney and lung. Isoform 1 is expressed in heart, kidney and
CC liver.
CC [19] INDUCTION: By E2F and p53 in apoptotic neurons.
CC [20] SIMILARITY: Contains 1 CARD domain.
CC [21] SIMILARITY: Contains 1 NB-ARC domain.
CC [22] SIMILARITY: Contains 13 WD repeats.
CC [23] CAUTION: Ref.7 sequence differs from that shown due to a
CC frameshift in position 109.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF013283; AAC51678.1; -
CC EMBL; AJ243003; CAB55579.1; -
CC EMBL; AJ243004; CAB55580.1; -
CC EMBL; AJ243005; CAB55581.1; -
CC EMBL; AJ243006; CAB55582.1; -
CC EMBL; AJ243007; CAB55583.1; -
CC EMBL; AJ243008; CAB55584.1; -
CC EMBL; AJ243009; CAB55585.1; -
CC EMBL; AJ243010; CAB55586.1; -
CC EMBL; AJ243011; CAB55587.1; -
CC EMBL; AJ243048; CAB55588.1; -
CC EMBL; AJ243107; CAB56462.1; -
CC EMBL; AF134397; AAD38344.1; -
CC EMBL; AF149794; AAD34016.1; -
CC EMBL; AB007873; BAA24843.2; ALT_INIT.
CC EMBL; AB103079; BAC77343.1; -
CC EMBL; AJ133643; CAB565085.1; -
CC EMBL; AJ133644; CAB565086.1; -
CC EMBL; AJ133645; CAB565087.1; -
CC EMBL; AF248734; AAK28401.1; ALT_FRAME.
CC PIR; T03818; T03818.
CC PDB; 1CL5; NMR; A=1-97.
CC PDB; 1CWM; NMR; A=1-97.
CC PDB; 1C15; X-ray; A=1-97.
CC PDB; 2YGS; X-ray; A=1-92.
CC PDB; 3YGS; X-ray; C=1-95.
CC Genew; HGNC:576; APAF1.
CC MIM; 602233; -
CC GO; GO:0005829; C:cytosol; TAS.
CC GO; GO:0008656; P:caspase activator activity; TAS.
CC GO; GO:0007399; P:neurogenesis; TAS.
CC GO; GO:0042981; P:regulation of apoptosis; TAS.
CC InterPro; IPR001315; CARD.
CC InterPro; IPR011029; DEATH_like.
CC InterPro; IPR002182; NB-ARC.
CC InterPro; IPR001680; WD40.
CC InterPro; IPR011046; WD40_like.
CC Pfam; PF00619; CARD; 1.
CC Pfam; PF00931; NB-ARC; 1.
CC Pfam; PF00400; WD40; 13.
CC PRINTS; PRO0320; GPROTEINRPT.
CC PRODOM; PD000018; WD40; 3.
CC PROSITE; PS50209; CARD; 1.
CC PROSITE; PS00678; WD_REPEATS_1; 4.
CC PROSITE; PS50082; WD_REPEATS_2; 9.
CC PROSITE; PS50294; WD_REPEATS_REGION; 1.

KW	3D-structure; Alternative splicing; Apoptosis; ATP-binding; Direct protein sequencing; Repeat; WD repeat.	
KW	Query Match	
	Best Local Similarity 100.0%; Score 6619; DB 1; Length 1248;	
	Matches 1248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MDKARNCLQHREALDKDIKTSYIMDMHISDGLFTISEEEKVRNEPTQOQRAAMLIKM	60
Db	1 MDKARNCLQHREALDKDIKTSYIMDMHISDGLFTISEEEKVRNEPTQOQRAAMLIKM	60
Qy	61 LKNDNSVVSFNALLHGYKDLAALLHDGIPVSSSGKDSVSGITSVRTVLCGGVP	120
Db	61 LKNDNSVVSFNALLHGYKDLAALLHDGIPVSSSGKDSVSGITSVRTVLCGGVP	120
Qy	121 QRPVVVTRKLVNAIQKSLKGEKPGWVTIHGMAGCKSVLAABAVRDSHSLGCFPG	180
Db	121 QRPVVVTRKLVNAIQKSLKGEKPGWVTIHGMAGCKSVLAABAVRDSHSLGCFPG	180
Qy	181 GVHWVSQKQKSGLLMLKQLNCTRLDQDSEFSQRLPLNIEBAKDRILMLRKHPRSL	240
Db	181 GVHWVSQKQKSGLLMLKQLNCTRLDQDSEFSQRLPLNIEBAKDRILMLRKHPRSL	240
Qy	241 ILDDVWDSWVLKAFDSQCIILLTRDKSVTSVWGPVVPVSSSLGKEGLIELSLFN	300
Db	241 ILDDVWDSWVLKAFDSQCIILLTRDKSVTSVWGPVVPVSSSLGKEGLIELSLFN	300
Qy	301 MKKADLPQAHSIIECKGSPVLSLIGALLRDPFNWYVYLKQIQKFKIRKSSSYD	360
Db	301 MKKADLPQAHSIIECKGSPVLSLIGALLRDPFNWYVYLKQIQKFKIRKSSSYD	360
Qy	361 YEALDEAMSISVEMLRDIDKYITDLSILQKDVKPTKVLCLMDMETEEVEDILQEFVN	420
Db	361 YEALDEAMSISVEMLRDIDKYITDLSILQKDVKPTKVLCLMDMETEEVEDILQEFVN	420
Qy	421 KSLFLCDRNGKSFYRIYDLQVDFTEKNCQLODLHKIITOFORYHQPHLTSPQSDC	480
Db	421 KSLFLCDRNGKSFYRIYDLQVDFTEKNCQLODLHKIITOFORYHQPHLTSPQSDC	480
Qy	481 MYWYNFLAYHNASAKWHKELCALMPSLDWIKAETELVGPFAHLIHFVYRHLDEKCAV	540
Db	481 MYWYNFLAYHNASAKWHKELCALMPSLDWIKAETELVGPFAHLIHFVYRHLDEKCAV	540
Qy	541 SENFQFLSLNGHLGRQFPFNIVOLGICEPETSSEVYQAKLQAKQEVNDGMLYLEWINK	600
Db	541 SENFQFLSLNGHLGRQFPFNIVOLGICEPETSSEVYQAKLQAKQEVNDGMLYLEWINK	600
Qy	601 KNITNLSRLVVRPHPTDAYVHACFSDGQRIASCADKTLQVFKATGKLEIKAHEDV	660
Db	601 KNITNLSRLVVRPHPTDAYVHACFSDGQRIASCADKTLQVFKATGKLEIKAHEDV	660
Qy	661 LCCAFSTDDRFTATCSVDKVKIWNMTGELVHTYDEHSEOVNCHFTNSSHHLLLATGS	720
Db	661 LCCAFSTDDRFTATCSVDKVKIWNMTGELVHTYDEHSEOVNCHFTNSSHHLLLATGS	720
Qy	721 SDCFLKLDLNLQKCRNTMFGHTNSVNRHCRSPDDKLLASCADGTLKLWATSANERKS	780
Db	721 SDCFLKLDLNLQKCRNTMFGHTNSVNRHCRSPDDKLLASCADGTLKLWATSANERKS	780
Qy	781 INVQKFFLNLEDQDMEVIVKCCSWASDAGRIWAAKNKIFLFIHTSGLLGEIHTGHH	840
Db	781 INVQKFFLNLEDQDMEVIVKCCSWASDAGRIWAAKNKIFLFIHTSGLLGEIHTGHH	840
Qy	841 STIOYCFSPQNHLLAVVALSOYCVELWNTDTSRVSQVADCRGHLVHWGVSPDGSFLTS	900
Db	841 STIOYCFSPQNHLLAVVALSOYCVELWNTDTSRVSQVADCRGHLVHWGVSPDGSFLTS	900
Qy	901 SDDQIRLWETKVKCNKSNVLMKQSDVVFQENVMVLAVDHIRLQNLINGRTGQIDYLT	960
Db	901 SDDQIRLWETKVKCNKSNVLMKQSDVVFQENVMVLAVDHIRLQNLINGRTGQIDYLT	960
Qy	961 EAQVSCCLSPHLQVIAFGDENGAIETLELVNRRIFQSRFQHKTVWHIQFTADEKTLIS	1020
Db	961 EAQVSCCLSPHLQVIAFGDENGAIETLELVNRRIFQSRFQHKTVWHIQFTADEKTLIS	1020

RESULT 2

APAF_MOUSE			
ID	APAF_MOUSE	STANDARD;	PRT; 1249 AA.
AC	O88879;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Apoptotic protease activating factor 1 (Apaf-1).		
GN	Name=Apaf1;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RC	STRAIN=Swiss Webster/NIH Swiss; TISSUE=Embryo;		
RX	MEDLINE=98424242; PubMed=9753320;		
RA	Cecconi F., Alvarez-Bolado G., Meyer B.I., Roth K.A., Gruss P.;		
RT	"Apaf1 (CED-4 homolog) regulates programmed cell death in mammalian development.";		
RT	Cell 94:727-737(1998).		
RL	[2]		
RN	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RC	TISSUE=Spleen;		
RX	MEDLINE=21063441; PubMed=1119689; DOI=10.1016/S0006-8993(00)02916-4;		
RA	Walke D.W., Morgan J.I.;		
RT	"A comparison of the expression and properties of Apaf-1 and Apaf-1L.";		
RL	Brain Res. 886:73-81(2000).		
CC	-!- FUNCTION: Oligomeric Apaf-1 mediates the cytochrome c-dependent autocatalytic activation of pro-caspase 9 (Apaf-3), leading to the activation of caspase-3 and apoptosis. This activation requires ATP (By similarity).		
CC	-!- SUBUNIT: Monomer. Oligomerizes upon binding of cytochrome c and dATP. Oligomeric Apaf-1 and pro-caspase-9 bind to each other via their respective NH2-terminal CARD domains and consecutively mature caspase-9 is released from the complex (By similarity). It may also interact with Bcl-XL.		
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).		
CC	-!- ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=2;		
CC	Name=1; Synonyms=Apaf-1L;		
CC	Isoid=O88879-1; Sequence=Displayed;		
CC	Note=Major isoform;		
CC	Name=2;		
CC	Isoid=O88879-2; Sequence=VSP_006763;		
CC	-!- TISSUE SPECIFICITY: Highly expressed in lung and spleen, weakly in brain and kidney and not detectable in liver.		
CC	-!- DEVELOPMENTAL STAGE: High levels in embryonic brain and liver from E11.5 to E17.5 day.		
CC	-!- SIMILARITY: Contains 1 CARD domain.		
CC	-!- SIMILARITY: Contains 1 NB-ARC domain.		
CC	-!- SIMILARITY: Contains 13 WD repeats.		

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF064071; AAC62458.1; --
HSP; Q9Y297; 1P22.
MG; MG1; 1306796; Apaf1.
CO; CO:0016505; F:apoptotic protease activator activity; IPI.
CO; CO:0007420; P:brain development; IMP.
CO; CO:0007275; P:development; IMP.
InterPro; IPR011029; CARD.
InterPro; IPR011029; DEATH_like.
InterPro; IPR002182; NB-ARC.
InterPro; IPR001680; WD40.
InterPro; IPR011046; WD40_like.
Pfam; PF00619; CARD; 1.
Pfam; PF00931; NB-ARC; 1.
Pfam; PF00400; WD40; 13.
PRINTS; PR00320; GPROTEINRPT.
ProDom; PD000018; WD40; 3.
SMART; SM00320; WD40; 13.
PROSITE; PS02029; CARD; 1.
PROSITE; PS00678; WD REPEATS; 1; 4.
PROSITE; PS00082; WD REPEATS; 2; 9.
PROSITE; PS0294; WD REPEATS_REGION; 1.
KW Alternative splicing; Apoptosis; ATP-binding; Repeat; WD repeat.
FT DOMAIN 1 90
FT NP_BIND 106 415
FT REPEAT 154 161
FT REPEAT 613 652
FT REPEAT 655 694
FT REPEAT 697 738
FT REPEAT 741 780
FT REPEAT 786 837
FT REPEAT 838 877
FT REPEAT 880 919
FT REPEAT 959 998
FT REPEAT 1001 1040
FT REPEAT 1042 1080
FT REPEAT 1083 1122
FT REPEAT 1125 1164
FT REPEAT 1176 1213
FT VARSPLIC 99 110
SQ SEQUENCE 1249 AA; 14088 MW; 53D91539CEFE6025C CRC64;
Query Match 88.7%; Score 5873.5; DB 1; Length 1249;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 1093; Conservative 79; Mismatches 76; Indels 1; Gaps 1;
Db 1 MDKARNCLLQHREALEKDKITSYIMDHMSIDGFLTISEEEKVNEPTQQRAAMLIKMI 60
Db 1 MDKARNCLLQHREALEKDKITSYIMDHMSIDGFLTISEEEKVNEPTQQRAAMLIKMI 60
Qy 61 LKNDNDVSFYNALLHEGYKDLAALHDIPIVWSSSGKDSVSGITSYVTVLCEGGVP 120
Db 61 LKNDNCAYISFYNALLHEGYKDLAALHDIPIVWSSSGKDSVSGITSYVTVLCEGGVP 120
Qy 121 QRPVVFVTRKLVNAIOOKLSKLGEPGWVTHGWAGCKSVLAABAVRHSILGECFPG 180
Db 121 QRPVVFVTRKLVNAIOOKLSKLGEPGWVTHGWAGCKSVLAABAVRHSILGECFPG 180
Qy 181 GHHVSVGKQDKSLGMLKQLNLTLDODESQRPLNIEAKDRILMLKHPRLSL 240
Db 181 GHHVSVGKQDKSLGMLKQLNLTLDODESQRPLNIEAKDRILMLKHPRLSL 240
Qy 241 ILDDVWDSVLKAFDSQCCILLTTRDKSVTDSVMGPKYVVPVSSGLGKGLKILSLFN 300
Db 241 ILDDVWDSVLKAFDSQCCILLTTRDKSVTDSVMGPKYVVPVSSGLGKGLKILSLFN 300

Qy 301 MKKADLPQAHSHIKCKGSPVLSLIGALLRDPNNRWYVYLKOLQNKQPKRIRKSSVD 360
Db 301 MKKEDLPAAHSHIKCKGSPVLSLIGALLRDPNNRWYVYLKOLQNKQPKRIRKSSVD 360
Qy 361 YEALDEAMSISVEMLRKEDIKDYVYDLSILQDKVYKPKVLCILWDMETEEVEDILOEFVN 420
Db 361 YEALDEAMSISVEMLRKEDIKDYVYDLSILQDKVYKPKVLCILWDMETEEVEDILOEFVN 420
Qy 421 KSLFLPCDRNGKSPRYVYLDLQVDFLTENCSQLODLHKKIITQRYVHPHTLSPQDEDC 480
Db 421 KSLFLPCNRNGKSPRYVYLDLQVDFLTENCSQLODLHKKIITQRYVHPHTLSPQDEDC 480
Qy 481 MYWYNFLAYHMASAKMVKELCALMESLQWIKAKTELGPAPHLIHEFVEYRHILDEKCAV 540
Db 481 MYWYNFLAYHMASAKMVKELCALMESLQWIKAKTELGPAPHLIHEFVEYRHILDEKCAV 540
Qy 541 SENFOEPLSLNGHLGROPPNIVOLGCEPETSEVVOAKLOAKQEVNDNGMLYLEWINK 600
Db 541 CENFOEPLSLNGHLGROPPNIVOLGCEPETSEVVOAKLOAKQEVNDNGMLYLEWINK 600
Qy 601 KNTNLSRLVVRPHTDAVHACGSEGGRIASCGADKTLQVKAETGEKLETKAHEDEV 660
Db 601 KNTNLSRLVVRPHTDAVHACGSEGGRIASCGADKTLQVKAETGEKLETKAHEDEV 660
Qy 661 LCCAFSTDDRIATCSVDKVKVWINSMTGELVHYVDHSEQVNCCHFTNSSHLLATGS 720
Db 661 LCCAFSTDDRIATCSVDKVKVWINSMTGELVHYVDHSEQVNCCHFTNSSHLLATGS 720
Qy 721 SDCLFLKLDLNQKECRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 780
Db 721 SDCLFLKLDLNQKECRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 780
Qy 781 INVQKQFLNLEDOEDMEVIVKCCSWADGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 840
Db 781 INVKRFLLSSEDPEDVEIVKCCSWADGDKIIWAAKNKIFLFDIHTSGLLGEIHTGHH 840
Qy 841 STIOYCDPSPONHLAVVALSOYCVELWNTDSKRVADCRGHLSSVHGVMPSPGSSFLTS 900
Db 841 STIOYCDPSPONHLAVVALSOYCVELWNTDSKRVADCRGHLSSVHGVMPSPGSSFLTS 900
Qy 901 SDDQIRLWETKKVCKNSAVMLKQEVVDVFOENVMVAVDHRILQINGRTGQIDYLT 960
Db 901 SDDQIRLWETKKVCKNSAVMLKQEVVDVFOENVMVAVDHRILQINGRTGQIDYLT 960
Qy 961 EAOVSCCLSPHLOYTARGDENGATLEILVNNRIFQSRFOHKKTVWHIOFTADEKTLIS 1020
Db 961 EAOVSCCLSPHLOYTARGDENGATLEILVNNRIFQSRFOHKKTVWHIOFTADEKTLIS 1020
Qy 1021 SDDABEQVNWQDKCIFLGRHOETVVDKFRLLKNSRLLSWSFDCTVKNWNIITGNKEVD 1080
Db 1021 SDDABEQVNWQDKCIFLGRHOETVVDKFRLLKNSRLLSWSFDCTVKNWNIITGNKEVD 1080
Qy 1081 FVCHQGTVLSCDISHDATKFSSTSDAKTAKIWSFDLLPLPLHGLHNGCVRCFAFVSDT 1140
Db 1081 FVCHQGTVLSCAISSDATKFSSTSDAKTAKIWSFDLLPLPLHGLHNGCVRCFAFVSDT 1140
Qy 1141 LIATGDDNGEIRIWNVNGELHLCAPLS-EEGAATHGCGWVTDLCFSPDGKMLISAGGYI 1199
Db 1141 LIATGDDNGEIRIWNVNGELHLCAPLS-EEGAATHGCGWVTDLCFSPDGKMLISAGGYI 1199
Qy 1200 KWNVVTGESSOTFYTGNTLNKKIHVSPPFTYTVVDNLGILYIQLTLE 1248
Db 1201 KWNVVTGESSOTFYTGNTLNKKIHVSPPFTYTVVDNLGILYIQLTLE 1248
RESULT 3
APAF_RAT ID APAF_RAT STANDARD; PRT; 1249 AA.
AC Q9Y297;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Apoptotic protease activating factor 1 (Apaf-1).
GN Name-Apaf1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=Sprague-Dawley;
RC Itoh T., Itoh A., Pleasure D.;
RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RN DEVELOPMENTAL REGULATION, AND INDUCTION BY BRAIN INJURY.
RX MEDLINE=21450943; PubMed=11567033;
RA Yakovlev A.G., Oca K., Wang G., Mavesyan V., Bao W.-L., Yoshihara K.,
RA Faden A.I.;
RT "Differential expression of apoptotic protease-activating factor-1 and
RT caspase-3 genes and susceptibility to apoptosis during brain
RT development and after traumatic brain injury.";
RL J. Neurosci. 21:7439-7446(2001).
CC -!- FUNCTION: Regulates programmed cell death; necessary for normal
CC brain development. Participates with caspase-9 (Apaf-3) in the
CC cytochrome c-dependent activation of caspase-3, leading to
CC apoptosis. This activation requires ATP (By similarity).
CC -!- SUBUNIT: Monomer. Oligomerizes upon binding of cytochrome c and
CC dATP. Oligomeric Apaf-1 and pro-caspase-9 bind to each other via
CC their respective NH2-terminal CARD domains (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- DEVELOPMENTAL STAGE: Highly expressed in brain cortex in embryos
CC (E17) and new-born rats up to day 7. Very low expression
CC thereafter.
CC -!- INDUCTION: By brain injury.
CC -!- SIMILARITY: Contains 1 CARD domain.
CC -!- SIMILARITY: Contains 1 NB-ARC domain.
CC -!- SIMILARITY: Contains 13 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF320222; AAG335067.1; -.
DR HSP; P16649; 1ERJ.
DR RGD; 620575; Apaf1.
DR InterPro; IPR001315; CARD.
DR InterPro; IPR011029; DEATH_like.
DR InterPro; IPR002182; NB-ARC.
DR InterPro; IPR004106; Peptidase_S9A_N.
DR InterPro; IPR001680; WD40.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF00619; CARD; 1.
DR Pfam; PF00931; NB-ARC; 1.
DR Pfam; PF00400; WD40; 12.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 3.
DR SMART; SM00320; WD40; 13.
DR PROSITE; PS50209; CARD; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 4.
DR PROSITE; PS50082; WD_REPEATS_2; 9.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Apoptosis; ATP-binding; Repeat; WD repeat.
FT DOMAIN 1 90
FT NP_BIND 106 415 NB-ARC.
FT REPEAT 154 161 ATP (potential).
FT REPEAT 613 652 WD 1.
FT REPEAT 655 694 WD 2.
FT REPEAT 697 738 WD 3.
FT REPEAT 741 780 WD 4.
FT REPEAT 796 837 WD 5.
FT REPEAT 838 877 WD 6.
FT REPEAT

FT REPEAT 880 919 WD 7.
FT REPEAT 959 998 WD 8.
FT REPEAT 1001 1040 WD 9.
FT REPEAT 1042 1080 WD 10.
FT REPEAT 1083 1122 WD 11.
FT REPEAT 1125 1164 WD 12.
FT REPEAT 1176 1213 WD 13.
SQ SEQUENCE 1249 AA; 141151 MW; 7B4A8116FAD008E9 CRC64;

Query Match 88.7%; Score 5871.5; DB 1; Length 1249;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1091; Conservative 82; Mismatches 75; Indels 1; Gaps 1;

QY 1 MDAKARNCLLOHREALEKDIKTSYIMDHIMISDGFITTESEBKVRNEPTQOQRAAMLIKMI 60
DB 1 MDAKARNCLLOHREALEKDIKTSYIMDHIMISDGFITTESEBKVRNEPTQOQRAAMLIKMI 60
QY 61 LKKNDNSVVSFNALLHGGYKDLAALLHDGIPVVSSSGSKDSVSGITSVYRTVLCGGVP 120
DB 61 LKNONYAISFNALLHGGYKDLAALLHSGPLVSSSSGKDTGGNTSFVRTVLCGGVP 120
QY 121 QRPVVFVTRKKLVNAIQOKLSKLAGEPGWVIHGMAGCGKSVLAEEAVRDHSLLEGCPFG 180
DB 121 QRPVVFVTRKKLVSAIQOKLWKLNGEPGWVIYGMAGCGKSVLAEEAVRDHSLLEGCPFG 180
QY 181 GVHWSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEEAKDRLRLMLRKHPRSL 240
DB 181 GVHWSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEEAKDRLRLMLRKHPRSL 240
QY 241 ILDDVMSWLVKAFDSQCOILLTTRDKSVTDSVNGPKYVVPVSSSLGKEGLEILSLFN 300
DB 241 ILDDVMPWLVKAFDQCOILLTTRDKSVTDSVNGPKYVVPVSSSLGKEGLEILSLFN 300
QY 301 MKADLPEQAHSHIIECKGSPVLSLICALRDPNRMWYVYLKOLQNKQFKRIRKSSYD 360
DB 301 MKKEDLPVSAHSIIECKGSPVLSLVGALLRDPNRMWYVYLQNLQNKQFKRIRKSSYD 360
QY 361 YEALDEAMSISVEMLREDIKDYTTLSILOKDVKVPVKVLCILWDMTEVEVEDILQEPVN 420
DB 361 YEALDEAMSISVEMLREDIKDYTTLSILOKDVKVPVKVLCVLDLETEVEVEDILQEPVN 420
QY 421 KSLIFCDNRNGSFYVYLDLQVDFLTNCSQLODLHKKIITQORVHOPHTLSPQSDC 480
DB 421 KSLIFCDNRNGSFYVYLDLQVDFLTNCSQLODLHKRMVITQORVHOPHTLSPQSDC 480
QY 481 MYWTFNFLAYHMASAKMKELCALMFLDWIKAKTELGVPAHLIHEFVYRHLDEKCAV 540
DB 481 MYWTFNFLAYHMASAGMKELCALMFLDWIKAKTELGVPAHLIHEFVYRHLDEKCAV 540
QY 541 SENFQSFSLNGLHLLGRQPPNIVOLGLCEPETSEVYQQAQKQAEVDNGLYLEWINK 600
DB 541 CENFQSFSLNGLHLLGRQPPNIVOLGLCEPETSEVYQQAQKQAEVDNGLYLEWINK 600
QY 601 KNIITNLSRLVVRPHTDAVYHACSEDQRIASCADKTLQVFKAEYTGKLEIKAEHDEV 660
DB 601 KTIKNSRLVVRPHTDAVYHACSEDQRIASCADKTLQVFKAEYTGKLEIKAEHDEV 660
QY 661 LCCAFSTDDEFIATCSVDKVKIWNMTGELVHTYVDEHSEQVNCCHFTNSHHLLATGS 720
DB 661 LCCAFSSDDSYIATCSVDKVKIWNMTGELVHTYVDEHSEQVNCCHFTNSHHLLATGS 720
QY 721 SDCFLKLWDLNQKECRNTFMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
DB 721 NDSFLKLWDLNQKECRNTFMFGHTNSVTHCRFSPDDELLASCADGTLKLDVRSANEKS 780
QY 781 INVQKPFNLNLEDPOEDMEVIVKCCSWADGARIWVAANKIFLDIHTSGLLGEIHTGHH 840
DB 781 INVKRFLLSSDEDPDEVEVIVKCCSWADGRIIVAAKNKVLDDIHTSGLLTIHTGHH 840
QY 841 STIQYCDFPSONHLAVVALSQYVELWNTDSRSKVADCRHLSVWHGVMPDGSFLTS 900
DB 841 STIQYCDFPSYDHLAVIALSQYVELWNIDSRVKVADCRHLSVWHGVMPDGSFLTA 900

```

QY 901 SDDQTRLMETKVKCNKSNVLMKQEVDDVFOENVMVLAVDHRLQLINGRTQCIDYLT 960
DB 901 SDDQTRVWETKVKCNKSAIIVLQEIIDVFOENVMVLAVDHRLQLIAKTQCIDYLP 960
QY 961 EAQVSCCLSPHLOVAFGDENGAIEILELVNRRIFQSRFOHKTVMHIOFTADEKTLIS 1020
DB 961 EAQVSCCLSPHLEVAFGDENGAIKIIELPNNRVFSSGIGHKKAVRHIOFTADGKTLIS 1020
QY 1021 SSDABEQVNMQLDKCIFLRHQETVKDFRLLKNSRLLSWSFGDTGVKVMNIIITGNKEKD 1080
DB 1021 SSDDSVIQVNMVQTEYVFLQAHQETVKDFRLLRDSRLLSWSFGDTGVKVMNVIITGRERD 1080
QY 1081 FVCHQGVLSCDISHDAYTKFSTGADTKAKIWSFDLPLHLRHLRGHNCVRCSAFSDVST 1140
DB 1081 FTCHQGVLSLCAISSDATKFSSTGADTKAKIWSFELPSPLHKLKGNHNCVRCSAFSLDGI 1140
QY 1141 LLATGDNGEIRIWNVSGELLHLHCAPLS-EEGAATHGGWVTDLCFSDGKMLISAGGYI 1199
DB 1141 LLATGDNGEIRIWNVSGQLLHLHCAPISIEEGTATHGGWVTDVCFSDRKLVSAGGYL 1200
QY 1200 KWNVNVYTGESSOTFYNGTNIKKTHVSPDKTYVTVDNLGILYILOTLE 1248
DB 1201 KWNVNVYTGESSQTFYNGTNIKKTHVSPDKTYVTVDNLGILYILOTLE 1249

RESULT 4
QSVI66
ID Q8VI66 PRELIMINARY; PRT; 1249 AA.
AC Q8VI66; 2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Apoptotic protease activating factor-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN SEQUENCE FROM N.A.
RP TISSUE=Cerebellum.
RC Cao G., Chen J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC EMBL; AF218388; AAL36935.1; -.
DR HSP; P16649; IERJ.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0002333; F:peptidase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0043829; P:defense response to pathogen; IEA.
DR GO; GO:0042981; P:regulation of apoptosis; IEA.
DR InterPro; IPR001315; CARD.
DR InterPro; IPR011029; DEATH_like.
DR InterPro; IPR000767; Disease_resist.
DR InterPro; IPR002182; NB-ARC.
DR InterPro; IPR001680; WD40.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF00619; CARD; 1.
DR Pfam; PF00931; NB-ARC; 1.
DR Pfam; PF00400; WD40; 12.
DR PRINTS; PR00364; DISEASERIST.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 3.
DR SMART; SM00320; WD40; 13.
DR PROSITE; PS02029; CARD; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 4.
DR PROSITE; PS00082; WD_REPEATS_2; 9.
DR PROSITE; PS02094; WD_REPEAT_REGION; 1.
KW Protease; Repeat; WD repeat.
SQ SEQUENCE 1249 AA; 141169 MW; 2B41B629599A9A38 CRC64;

```

Query Match 88.6%; Score 5863.5; DB 2; Length 1249;
 Best Local Similarity 87.2%; Pred. No. 0;

```

Matches 1089; Conservative 84; Mismatches 75; Indels 1; Gaps 1;
QY 1 MDKARNCLLOHREALEKDIKTSYIMDMHISDGFITISEEEKVRNEPTQOQAAMLIKMI 60
DB 1 MDKARNCLLOHREALEKDIKTSYIMDMHISDGFITISEEEKVRNEPTQOQAAMLIKMI 60
QY 61 LKKNDSVSVSYNALLHEGYKDLAALLHDGIPVSSSSGKDSVSGITSYVVRTVLCGGVP 120
DB 61 LKKNDSVSVSYNALLHEGYKDLAALLHDGIPVSSSSGKDSVSGITSYVVRTVLCGGVP 120
QY 121 QRPVVFTRKKLVNAIQKLSKLGEPGWITIHGMAGCGKSVLAABAVRDHSLLECGPFG 180
DB 121 QRPVVFTRKKLVNAIQKLSKLGEPGWITIHGMAGCGKSVLAABAVRDHSLLECGPFG 180
QY 181 GVHWVSKGDKSGLMKLQNLCTRLDQDESFSORLPLNTEAKDRLLIIMLRKPRSL 240
DB 181 GVHWVSKGDKSGLMKLQNLCTRLDQDESFSORLPLNTEAKDRLLIIMLRKPRSL 240
QY 241 ILDDVWSWVLCAPDSQCQILLTTRDKSVTSVMGPKYVVPVSSSGKEKGLHLSLFW 300
DB 241 ILDDVWSWVLCAPDSQCQILLTTRDKSVTSVMGPKYVVPVSSSGKEKGLHLSLFW 300
QY 301 MKKADLPQAHSIIECKGSPVWSLIGALLRDPFNWYLLKOLQKQPKRIRKSSSYD 360
DB 301 MKKADLPQAHSIIECKGSPVWSLIGALLRDPFNWYLLKOLQKQPKRIRKSSSYD 360
QY 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKDVPTKVLCLWDMETEVEEDILOEPV 420
DB 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKDVPTKVLCLWDMETEVEEDILOEPV 420
QY 421 KSLFLCNRNGKSFYLLHDLQVDFLTEKNSQLODLHKIITOFQYHOPHTLSPQDQDC 480
DB 421 KSLFLCNRNGKSFYLLHDLQVDFLTEKNSQLODLHKIITOFQYHOPHTLSPQDQDC 480
QY 481 MYWNYFLAYHMAKMKELCALMFSLDWIKAKTELGVPAHLIHEFVYRHLDEKCAV 540
DB 481 MYWNYFLAYHMAKMKELCALMFSLDWIKAKTELGVPAHLIHEFVYRHLDEKCAV 540
QY 541 SENFOEFLSLNGHLGRQPPNIVOLGCEPETSEVVOAKLOAKOEVDNGMYLEWINK 600
DB 541 SENFOEFLSLNGHLGRQPPNIVOLGCEPETSEVVOAKLOAKOEVDNGMYLEWINK 600
QY 601 KNITNLRLVVRPHTDANVHACFSDEGORIASCADTKLOVFAETGEKLEIKAHEDV 660
DB 601 KNITNLRLVVRPHTDANVHACFSDEGORIASCADTKLOVFAETGEKLEIKAHEDV 660
QY 661 LCCAFSTDDRFATCSVDKVKIKNMTCGLVHTYDSEOVNCHFTNSHHLLIATGS 720
DB 661 LCCAFSTDDRFATCSVDKVKIKNMTCGLVHTYDSEOVNCHFTNSHHLLIATGS 720
QY 721 SDCFLKLDLNLQKECNTWFGHTNSVNHCRFSPDDKLLASCADGTLKLWDATSANERKS 780
DB 721 SDCFLKLDLNLQKECNTWFGHTNSVNHCRFSPDDKLLASCADGTLKLWDATSANERKS 780
QY 781 INVKOFFIANLEDPOEDMEVIVKCCSWASDAGARIMVAAKNKIIFLFDIHTSGLLGEIHTGHH 840
DB 781 INVKOFFIANLEDPOEDMEVIVKCCSWASDAGARIMVAAKNKIIFLFDIHTSGLLGEIHTGHH 840
QY 841 STIQCYDFSQNHAVVALSQQVCELMNTDSRKSVDACRGLHSVHGVNMFSPDGSFLTS 900
DB 841 STIQCYDFSQNHAVVALSQQVCELMNTDSRKSVDACRGLHSVHGVNMFSPDGSFLTS 900
QY 901 SDDQTRLMETKVKCNKSNVLMKQEVDDVFOENVMVLAVDHRLQLINGRTQCIDYLT 960
DB 901 SDDQTRLMETKVKCNKSNVLMKQEVDDVFOENVMVLAVDHRLQLIAKTQCIDYLP 960
QY 961 EAQVSCCLSPHLOVAFGDENGAIEILELVNRRIFQSRFOHKTVMHIOFTADEKTLIS 1020
DB 961 EAQVSCCLSPHLEVAFGDENGAIKIIELPNNRVFSSGIGHKKAVRHIOFTADGKTLIS 1020
QY 1021 SSDDABEQVNMQLDKCIFLRHQETVKDFRLLKNSRLLSWSFGDTGVKVMNIIITGNKEKD 1080
DB 1021 SSDDABEQVNMQLDKCIFLRHQETVKDFRLLKNSRLLSWSFGDTGVKVMNIIITGNKEKD 1080
QY 1021 SSDDSVIQVNMVQTEYVFLQAHQETVKDFRLLRDSRLLSWSFGDTGVKVMNVIITGRERD 1080
DB 1021 SSDDSVIQVNMVQTEYVFLQAHQETVKDFRLLRDSRLLSWSFGDTGVKVMNVIITGRERD 1080

```

Qy 1081 FVCHQGVTLSCDISHDATKFSSTGADTKATKWSFDLLPLHLRHNGCVRCSAFSDVST 1140
 Db 1081 FTCHQGVTLSCAISDADTKFSSTGADTKATKWSFELRSPHLRHNGSCVRCSAFSLDGI 1140
 Qy 1141 LLATGDDNGEIRIWNVSGELHLHCAPLS-BEGAATHGCGWTDLCFSPDGGKWLISAGGYI 1199
 Db 1141 LLATGDDNGEIRIWNVSDGQLHLHCAPISIEGATHGCGWTDVCFSPDRKWLVSAGYL 1200
 Qy 1200 KWNVVTGESSQTFTYTGNTLNKKIHVSDFDKTYTYVDNLGLYILQTL 1248
 Db 1201 KWNVVTGESSQTFTYTGNTLNKKIHVSDFDKTYTYVDNLGLYILQVLE 1249

RESULT 5
 Q6GNU6 PRELIMINARY; PRT; 1248 AA.
 AC Q6GNU6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OK NCBI_TaxID=8355;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative."
 RL Dev. Dyn. 225:384-391(2002).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S., Strausberg R.;
 RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 11 WD repeats.
 DR EMBL; BC073405; AAH73405.1;
 DR InterPro; IPR01029; DEATH_like.
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR002182; NB-ARC.
 DR InterPro; IPR01680; WD40.
 DR InterPro; IPR011046; WD40_like.

DR Pfam; PF00619; CARD; 1.
 DR Pfam; PF00931; NB-ARC; 1.
 DR Pfam; PF00400; WD40; 11.
 DR PRINTS; PRO0364; DISEASERSIST.
 DR PRINTS; PRO0320; GPROTEINBRPT.
 DR SMART; SM00018; WD40; 2.
 DR SMART; SM00320; WD40; 13.
 DR PROSITE; PS02029; CARD; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 3.
 DR PROSITE; PS00082; WD_REPEATS_2; 9.
 DR PROSITE; PS02094; WD_REPEATS_REGION; 1.
 KW Hypothetical protein; Repeat; WD repeat.
 SQ SEQUENCE 1248 AA; 141574 MW; 822F7CEAC4EE675D CRC64;

Query Match 61.2%; Score 4054; DB 2; Length 1248;
 Best Local Similarity 60.5%; Pred. No. 3e-227;
 Matches 757; Conservative 198; Mismatches 289; Indels 8; Gaps 4;

Qy 1 MDKARNCLLOHREALEKDIKTSYIMDHMSDGLFTTISESEKVRNEPTQOORAAWLKMI 60
 Db 1 MDEKARSLLOHREALEKDIKTSYIMDHMSDGLFTTISESEKVRNEPTQOORAAWLKMI 60
 Qy 61 LKQNDSDYVSFYNNALLHEGYKDLAALLHLDGIPVVSSSSGKSDSVSGITSYVTVLCEGVP 120
 Db 61 LKQNDSDYVSFYNNALLHEGYKDLAALLHLDGIPVVSSSSGKSDSVSGITSYVTVLCEGVP 120
 Qy 121 QRPVVFVTRKLVNAIQOKLSKLKGEFGWVTIHMAGCGKSVLAEEAVRDSHSLLEGCFPG 180
 Db 121 QRPVVFVTRKLVNAIQOKLSKLKGEFGWVTIHMAGCGKSVLAEEAVRDSHSLLEGCFPG 180
 Qy 181 GVHWVSVGKDGKGLLMLKQLNCTRLDQDSEFSORLPINIEEAKDRILMLRKHPSLL 240
 Db 181 GVHWVSVGKDGKGLLMLKQLNCTRLDQDSEFSORLPINIEEAKDRILMLRKHPSLL 240
 Qy 241 ILDDVDSDWLVKAFDSOCQILTLTRDKSVTDSVVMGPKVVPVSVSSSLGKEGLEILSLFVN 300
 Db 241 ILDDVDSDWLVKAFDSOCQILTLTRDKSVTDSVVMGPKVVPVSVSSSLGKEGLEILSLFVN 300
 Qy 301 MKKADLPQEAHSIIKECKSPVLSLIGALLRDPNRPWEYLLKOLNQKPKRIRKSSSYD 360
 Db 301 MKKADLPQEAHSIIKECKSPVLSLIGALLRDPNRPWEYLLKOLNQKPKRIRKSSSYD 360
 Qy 361 YEALDEAMSTISVEMLRDIDYDLSILOKDVKVPKVLICILWDMTEVEVEDILOEFVN 420
 Db 361 YEALDEAMSTISVEMLRDIDYDLSILOKDVKVPKVLICILWDMTEVEVEDILOEFVN 420
 Qy 421 KSLILFCDRNGKSPRYLHLDQVDFLTETKNCSQLDLHKIITQFQYHQPHTLSPDOEDC 480
 Db 421 KSLILFCDRNGKSPRYLHLDQVDFLTETKNCSQLDLHKIITQFQYHQPHTLSPDOEDC 480
 Qy 481 MYWYNFLAYHMAKMHKEICALMFLSDWTAKTELVPAPHLIHEFVYRHLDEKCAV 540
 Db 481 MYWYNFLAYHMAKMHKEICALMFLSDWTAKTELVPAPHLIHEFVYRHLDEKCAV 540
 Qy 541 SENPQFELSLNGHLLGRQFPFNIVQLGCEPTESEVYQOAKLOAKQOEVDNGMLYLEWINK 600
 Db 541 SENPQFELSLNGHLLGRQFPFNIVQLGCEPTESEVYQOAKLOAKQOEVDNGMLYLEWINK 600
 Qy 601 KNTNLSRLVVRPHDTAVYHACFSDGQRIASCADKTLQVFKAEKGKLEIKAEHDEV 660
 Db 601 KNTNLSRLVVRPHDTAVYHACFSDGQRIASCADKTLQVFKAEKGKLEIKAEHDEV 660
 Qy 661 LCCAFSTDDRFIATCSVDKVKIWNMTGLVHTYDEHSEQVNCCHFTNSHHLLLATGS 720
 Db 661 LCCAFSTDDRFIATCSVDKVKIWNMTGLVHTYDEHSEQVNCCHFTNSHHLLLATGS 720
 Qy 721 SDCLPKLWLNQKRCRTMFGHTNSVNHCRFSPDDKLLASCSDGTLKLWDATSANERKS 780
 Db 721 SDCLPKLWLNQKRCRTMFGHTNSVNHCRFSPDDKLLASCSDGTLKLWDATSANERKS 780
 Qy 781 INVQKQFPLENLEDPQEDMEVIVKCSWSADGARIMVAANKIFLFIHTSGILGHIHTGH 840
 Db 781 INVQKQFPLENLEDPQEDMEVIVKCSWSADGARIMVAANKIFLFIHTSGILGHIHTGH 840
 Qy 838 IEVAKLFENEDESQF--EVLKKCCAWNSDGRINMTVTRNFLICIFDSTSCDLLSQLKACHQ 838
 Db 838 IEVAKLFENEDESQF--EVLKKCCAWNSDGRINMTVTRNFLICIFDSTSCDLLSQLKACHQ 838

QY 841 STIQYCDSPQNHLLAVVALSQYCVELWNTDSRSKVADCRGHLWSVHGYMSPDGSFSLTS 900
 Db 839 --ILYCDFCTTQNIQVALALSYYVQLWMDISSTKIAFNAHLSWHVCVKSPKSSSFLTS 896
 QY 901 SDQTRIMWETKVKCKSAVNLKQEVVFOENEVNVLAADVHRRQLQINGRTQIDYLT 960
 Db 897 SDDQTVKLVWETSNVSKPSATNLKREFDVSPNGEETLVLATSKDDCILLINGMTGETLSQI 956
 QY 961 EAG--VSCCLSPHLOVIAFGDENGAGIETLELVNRRIFQSFQHKTKVWHIOFTADEKT 1017
 Db 957 NTQDKCVTCCLTNDYQALAGDEGKVKVIDVSRGEILCKLGDHSATVQHCQFTADGKH 1016
 QY 1018 L1SSSDARIQVNWQDKCIFRCHQETVKQFRLLKNSRLLSNSFDGTVKWNITGNK 1077
 Db 1017 LVSSSDSTIRVWSLASGSELHKGHPKFKILQSQVFSFDGTVKWDLTGKL 1076
 QY 1078 EKDFVCHQGTULSCDISHATKSSSTADTKATKINSFDLLPLHLRCHNGCVRCGAFSV 1137
 Db 1077 RFEFICHSETVLSCDISPDSSTKSTADKSKINSLDMSTLLHNLGHQSCVRCRCFSW 1136
 QY 1138 DSTLLATGDNGEIRIMVNSVNGELLHLCAPLSEGA-ATHGQWVTDLCFGPDGKMLISAG 1196
 Db 1137 DNKYLATGDNGKIMVNSVQNGELLKQCCDISVNNENSLHDGWTDLHFSFNSKLIVSSG 1196
 QY 1197 GYIKWNVVTESSQTFYNTGNLKKIHVSPDKTVVVDNLGILYILOTLE 1248
 Db 1197 ANYKXWVDVDTGACKQTFYTVGTNLKSHLCPTFTFTVIDNLGILYIYKLTIE 1248
 RESULT 6
 APAF BRARE STANDARD; PRT; 1261 AA.
 AC Q919H8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Apoptotic protease activating factor 1 (Apaf-1).
 GN Name-apaf1
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OK NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20373792; PubMed=10917738;
 RA Inohara N., Nunez G.;
 RT "Genes with homology to mammalian apoptosis regulators identified in
 Zebrafish.";
 RL Cell Death Differ. 7:509-510(2000).
 CC -!- FUNCTION: Oligomeric Apaf-1 mediates the cytochrome c-dependent
 CC autocatalytic activation of pro-caspase 9 (Apaf-3), leading to the
 CC activation of caspase-3 and apoptosis. This activation requires
 CC ATP (By similarity).
 CC -!- SUBUNIT: Monomer. Oligomerizes upon binding of cytochrome c and
 CC DATP (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Contains 1 CARD domain.
 CC -!- SIMILARITY: Contains 1 NB-ARC domain.
 CC -!- SIMILARITY: Contains 13 WD repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF251502; RAJ67189.1; -.
 CC HSPSP; P16649; 1ERK.
 CC ZFIN; ZDB-GENE-0000616-4; apaf1.
 DR

InterPro; IPR001315; CARD.
 InterPro; IPR002182; NB-ARC.
 InterPro; IPR001680; WD40.
 InterPro; IPR011046; WD40_like.
 Pfam; PF00619; CARD; 1.
 Pfam; PF00931; NB-ARC; 1.
 Pfam; PF00400; WD40; 11.
 PRINTS; PR00320; GPROTEINRPT.
 ProDom; PD000018; WD40; 2.
 SMART; SM00114; CARD; 1.
 SMART; SM00320; WD40; 14.
 PROSITE; PS50209; CARD; 1.
 PROSITE; PS50678; WD_REPEATS_1; 2.
 PROSITE; PS50082; WD_REPEATS_2; 7.
 PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KMW Apoptosis; ATP-binding; Repeat; WD repeat.
 FT DOMAIN 1 90
 FT NP_BIND 106 417 NB-ARC.
 FT REPEAT 154 161 ATP (potential).
 FT REPEAT 615 654 WD 1.
 FT REPEAT 657 696 WD 2.
 FT REPEAT 700 743 WD 3.
 FT REPEAT 746 785 WD 4.
 FT REPEAT 798 836 WD 5.
 FT REPEAT 840 879 WD 6.
 FT REPEAT 882 921 WD 7.
 FT REPEAT 964 1003 WD 8.
 FT REPEAT 1006 1045 WD 9.
 FT REPEAT 1047 1088 WD 10.
 FT REPEAT 1091 1130 WD 11.
 FT REPEAT 1133 1172 WD 12.
 FT REPEAT 1184 1223 WD 13.
 SQ SEQUENCE 1261 AA; 142933 MW; F80CC9CB41764F1C CRC64;
 Query Match 55.7%; Score 3683.5; DB 1; Length 1261;
 Best Local Similarity 54.8%; Pred No. 1.2e-205;
 Matches 691; Conservative 244; Mismatches 310; Indels 15; Gaps 7;
 QY 1 WDARVNCILQHREALEKDKTSYMDHMSIDGELTISEBEKVNEPTCOORAMLIKMI 60
 Db 1 MEERARELLRSKATLEQDIKASYLMDHMSIDGVLTNDEAKVLKATKKEQVALLET 60
 QY 61 LKQNDSTVFYVALLHEGYKDLAALLHGDGIPVSSSGKDSVSGTSTVTVLCEGVP 120
 Db 61 LRKONRAYISFYNALITRESYGDLSLHSLDPLSPGEEKSFADGVSVPVQALLSGVGP 120
 QY 121 QRPVVFVTRKLVNAIOOKLSKLKGEPGVTTHGMAGCKSVLAABAVRDHSLLEGCPFG 180
 Db 121 QRPVVFVSRPPLNLIREMLYQLRDTPGWTVFGMAGSKSVMAAEVVRDRSLIKECPD 180
 QY 181 GVHVSVVGKQDKSGLMLQLNLCRLDQDES--FSQRLPLNIEAKDRILRLMRKPRS 238
 Db 181 GVHWSVGQCEERADLLVRMQSLCFRLEQCQSSDTSQRPPTVBEAKERLFLMLRRFPRS 240
 QY 239 LLILDDVWDSWMLKAFDSQCQILLATTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSF 298
 Db 241 LLILDDVWDSWSSLSRFDICRVLLTTRNRLTDSVGVRYVEVPVNGLDEKALBILALY 300
 QY 299 VMKKADLPQAHSIKECKGSPVLSLIGALLRDFNRWEYVLKOLONKQKFRIRKSSS 358
 Db 301 VNGRMHKLPEQARSIVSECKGSPVLSLIGALLRDFNRWSSYVLRQLQKQKFRIRKSSS 360
 QY 359 YDVALDEAMSISVEMLRDIDKYVTDLSILQDKVPTKVLCTILMDMETEEVDILQEF 418
 Db 361 YDVALQDAMDASLOVLEAEHQELYRDLISVMKDIKVPKVLVNLGLEBEEVDVLQEF 420
 QY 419 VNKSLIFCDNGKSFYLYLHDIQVDFLEKCSYLQDLHKKLTQORVHOPHTLSPOE 478
 Db 421 VNKSLIFCDNCQPYRYLHDIQVDFLEKCSYLQDLHKKLTQORVHOPHTLSPOE 480
 QY 479 DCMWYNYFLAYHASAKMKEKLCALMFSLDWIKATKTELVPALHIEFVEYRHLDEKDC 538
 Db 481 DLSYWYQFIPIYHMAKAGLSKELYSLMFSLDWKVKAKIMGSAHLINDYVEYGEILDKENS 540


```
Qy 539 AVSENFQFSLNGHLLQRPFPNIVQLGCEPSETSEVYQAKLQAKQEVNQMGLYLEWI 598
Db 541 EVRVQFQFSLNGHLLQRPFPNIVQLGCEPSETSEVYQAKLQAKQEVNQMGLYLEWI 600
Qy 599 NKNITNLRLVVRPHTDAVYHACSEDPORIASCGADKTLQVKAETGEKLEIKAHED 658
Db 601 KNTEEGSLRVLVMPHQVYIACFKDSKIASCGASKALURVFASTSGEKLLBELQAHEE 660
Qy 659 EVLCAFDSTDRFIATCSVDKVKIWNMTGELVHYD-EHSEQVNCCHFTNSHHLLLA 717
Db 661 DVLCAFDSPDRHIATCASDRKVLNVRGVLIREFEVEHEEQINHCQFTNTGRVLLA 720
Qy 718 TGSSDCP--LKLNDLNQKCEKNTWPGHTNSVNHCRFPSPDDKLLASCSDAGTFLKWDATSA 775
Db 721 TCSNDKFTNTLWPNKKTQNTWPGHMEPVNHCFCSPNDLYLATSSDGSGLKFEVSSA 780
Qy 776 NERLSNVYQFNLNEDPOEDMEVIVKCCSASDAGARLWAAKNKIFLFIHTSGLGEI 835
Db 781 NEWKSLDVSFPF---PESDEBKAMVKCSTWSADGSIICARNTVFVDFVETSDLLKL 837
Qy 836 HTGHSHTIQCDFSPQNHVALVVALSQCVELWNTDSRSKVADCRGHLVWVHGVFSPDGS 895
Db 838 KTSRLSTIQFCHACPNSSLLAVALSHYTVELWNFESSKKKAECSGHLVWVHGVFSPDGS 897
Qy 896 SFLTSSDDQTLRLWETKVKCNKSNVAMLKQEVVDVFOENVMVLAVDHRLRLQLINGRTGQ 955
Db 898 LLLSSDDQTLRLWETDRVHTSSAVALKRDVLSHSDATIIAPDSSNRQLVLSGSGTGA 957
Qy 956 IDYLTE---AQVSCCLSPHLQYAFGDENGAEILELVNRRIFQSRFQHKKTWHIOFT 1012
Db 958 VVLESEELSSRCSIRCSIRNAAFVALGSEDGTQVIEVPSSKASVKLSGGHTKTVHHCOFT 1017
Qy 1013 ADEKTLSSDDAEIQVNWOLDKIFLRGQETVKDFRLKNS---RLLSWSFDGTVKV 1069
Db 1018 DDCIILLTSSSDSIRVWKTGECVLOGHMEPVKPHLLSSSSPHLFSWSFDGTVKV 1077
Qy 1070 WNIITGNKEKDFVCHQGTVLSGDSHDKTSSTSDAKTAKIWSFDLLPLHLRLHGNGC 1129
Db 1078 WDLTRGQMLQDLVCHGAVLSCDVSSDGRLEFATTSAKTAKVMSSASWMLFLEGHKDC 1137
Qy 1130 VRCSAFSDVSTLLATGDDNGEIRIWNVNSGELLHLCAPLSEGA-ATHGGVNTDLCSFD 1188
Db 1138 VRSCRFSDNKRLLATGDDNGEIRLWSMDGALLKICPRDXTKDSMNSVHAGVNTDLHSPD 1197
Qy 1189 GKMLISAGYIKWNVVTVGSSQFTYTGNTLKKIHSVDPFKTYVDNLGILYIOTLE 1248
Db 1198 NRVLVSTAGYIKWNSVESGEALQFYTMGGNKKIHSVSPDFSTITVDSIGILYILKLE 1257

RESULT 7
Q80VR5 PRELIMINARY; PRT; 258 AA.
AC Q80VR5;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Similar to apoptotic protease activating factor 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=2238257; PubMed=12477932;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
```

```
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fabey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.B., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RC SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045144; AAH45144.1; -
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0042981; P:regulation of apoptosis; IEA.
DR InterPro; IPR001315; CARD.
DR InterPro; IPR011029; DEATH like.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00619; CARD; 1.
DR Pfam; PF00931; NB-ARC; 1.
DR PROSITE; PS50209; CARD; 1.
KW Protease.
SQ SEQUENCE 258 AA; 28808 MW; E1C16DEAD1D370A9 CRC64;
Query Match 16.1%; Score 1066; DB 2; Length 258;
Best Local Similarity 86.1%; Pred. No. 3.6e-54;
Matches 204; Conservative 18; Mismatches 15; Indels 0; Gaps 0;
Qy 1 MDAKARNCLLQHREALEKDKITSYIMDHMISDGLFTITSEEEKVNEPFTQQRAAMLIKMI 60
Db 1 MDAKARNCLLQHREALEKDKITSYIMDHMISDGLFTITSEEEKVNEPFTQQRAAMLIKMI 60
Qy 61 LKKDNDVSVFYNALHHEGYKDLAALLHDGTPVYVSSSGKDSVSGITSYVTVLCEGVP 120
Db 61 LNKDNCAYISFYNALHHEGYKDLAALLQSLGPLVSSSGKDTDGGTISFVTVLCEGVP 120
Qy 121 QRPVVFVTRKLVNAIOOKLSKLGEGPWTTHGMAGCGKSVLAEEAVRDRHSLLEGCPFG 180
Db 121 QRPVVFVTRKLVNAIOOKLSKLGEGPWTTHGMAGCGKSVLAEEAVRDRHSLLEGCPFG 180
Qy 181 GVHWVSVGKQKSGLLMKLQNLCTRLDQDSFSQRLPLNTEBEAKDRILMLRKHPR 237
Db 181 GVHWVSGIKQKSGLLMKLQNLCTRLDQDSFSQRLPLNTEBEAKDRILMLRKHPR 237

RESULT 8
Q820R1 PRELIMINARY; PRT; 1227 AA.
ID Q820R1;
AC Q820R1;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE WD-40 repeat protein.
GN OrderedLocustNames=alr0029;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=2159840; PubMed=11759840;
RX Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
```



```
RP SEQUENCE FROM N.A.
RX STRAIN=PCC 7421;
RA MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Teuchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003).
CC -!- SIMILARITY: Contains 14 WD repeats.
DR EMBL: AP006583; BAC92292.1; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0006915; P:apoptosis; IEA.
DR GO: GO:0042829; P:defense response to pathogen; IEA.
DR InterPro: IPR000767; Disease_resist.
DR InterPro: IPR002182; NB-ARC.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00931; NB-ARC; 1.
DR Pfam: PF00400; WD40; 14.
DR PRINTS: PR00364; DISEASERESIST.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PRODOM: PD000018; WD40; 13.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_7.
DR PROSITE: PS00682; WD_REPEATS_2; 14.
DR PROSITE: PS02394; WD_REPEATS_REGION; 1.
DR Complete proteome; Repeat; WD repeat.
SQ SEQUENCE 1184 AA; 128483 MW; 228C84D39C5A64F7 CRC64;

Query Match 10.5%; Score 693; DB 2; Length 1184;
Best local similarity 23.8%; Pred. No. 1.8e-31;
Matches 275; Conservative 172; Mismatches 469; Indels 238; Gaps 36;

QY 150 VTIHMGAGCGSVLAEEAVRDHSLLLEGCPGGVHVHVGKQKSLGKMLKLNLCRLDOD 209
DB 138 VAVLGGGIGTKALAVKLAROS---EAHFERLI-WRSL--RNAPPLADLAEILVFAGDE 191
QY 210 EGSFQRLPLNIEAKDLRIILMKRPHRSLLIIDDVWDSWLK-----AFDSQCQ---- 259
DB 192 Q--APALPTGIEGR--ILRLLECLRRNRCCLLVLDNA--ESLLRGEQAGTVREGCEGYGE 245
QY 260 -----ILLTRDSKVTDSDVMGPKYVVPVRESSLGKKEGLEILSLFVNMKADL 306
DB 246 LLRRVGEVPHASCLVLSREK-----PAEVALEGSSLPVRSRLGGL--QESEGEV 295
QY 307 PEQAHSI-----IKEC-KGSPLVSLIGALLRD-FPNRWEYILKQ-----L 345
DB 296 ILQAKGLRGADERRKLVCEYRGNFLALMIISTISIRELFDGQIREFLAQDTAVFNGIANL 355
QY 346 QNKQFKRIKRS--SSVDYEALDEAMSIVEMLRDIDKYDTLDSILQKDVKVPKVLCLILW 404
DB 356 LOQFDRLEAEKQMLWLAIHREPVVAQLRENV-----VPA----- 393
QY 405 DMETEEVEDILOEFVNKSLIFCDRNGKSFYVYHLDVDFTEKNCQLODLHKLIITOP 464
DB 394 -VSAPKLEALELSRLSLIERGTGTGQVPMVEYTFEHFVERIC--DEVHKGITGLF 449
QY 465 QRYHQPTLSPDQEDCMWYNFLAYHWSAKMHELKALMFLSLDWIKAKTELVGPAHLIH 524
DB 450 RSH-----ALLLAQAK-----YIIRAAQSRLILKPVID 477
QY 525 EFV----EYRHILBEKCAVSENFQBELSINGHLIGROFPFNNIVOLGCRPETSSEYVQQA 580
DB 478 ELLVRLGSOAH--LEORLAALVAQORDEAPLOFGYVGGNTL--NMLATLQTLRHWDFSLHA 535
QY 581 KLQAKQEVNDGMLYLEWINKNI-----TNLSRLVVRPHTDVYHACFSEDEQRIASCAD 636
DB 536 VMQA-----YLQEVNLYGVNFHGTDLARCVAQNFQGVGVFVAFSPDGEQIAGVDN 586
QY 637 KTLQVFKATGEKLEIKAEHEDEVLCCAFSTDRFIATCSVDKVKVKNWSMTGELVHTYD 696
DB 587 SEIRLWRAADQQQLSCQGHGTWVCVAFAPNGQTFASASQDGTGTVKLWDARIQCCLATLR 646
```

```
QY 697 EHSEQVNCCHFTNSSHLLIATGSSDCFLKLDLWLNQKCECRNTMFGHTNSVNHCRFSPDDK 756
DB 647 GHIGWVRSAAAPADGGS--LLASAGQDSTVKLWDAATGRCLATLQHTGGVHVSFAFAPDGS 704
QY 757 LLASCSADGTLKLDWATSANERKSINVKQFFLNLEDPOEDMEVIVKCCSWASDAGARIMVA 816
DB 705 LLASAGQDSTVKLWDAATG-----RCLA----- 727
QY 817 AKNKIFLFDIHTSGLLGEIHTGHHSIOYCDSPQNHVAVLVALSQYCVELWNTDSRSKVA 876
DB 728 -----TLQGHTEPIRSVWFSPDGHRLASASHDRTVKLNFPATGRCILA 769
QY 877 DCRGHLGSHVHWGWFSPDGSFLTSSDDQITRLWETKVKCKNSAVMLKQEVVVVFQ----- 931
DB 770 TLAGHGDWVSVAFAFADGRSLATGSLDRVRLWEI--ITGQCLKTLOEHTDQVFSIAPH 827
QY 932 -----ENVMVLAVDHRRLQINGRTGQIDYILTEAQVSCCLSPHQIYIARGDE 981
DB 828 QGHTLASGSPQTQVKLWDTESGQCLRTQGGKI-----VTVLAVAFSPHGQTLVSGSD 879
QY 982 NGATEILELVNRRIFQSRFOHKKTWHIOFTADEKTLISSDDDBAEIQVNWQLDKCI-FL 1040
DB 880 DRLVRLMDVRTGECTVRLRGLRGVTVVAVAPDGRKTLASAGADLSVKIMDALSGQCCLRTL 939
QY 1041 RGHQETVNDPELLKNGRLL-SWSFDGTVKVNIITGNKEKDFVCHQGTVLSCDISHDATK 1099
DB 940 REHTGSRIRSVAFAPDGRLLASGSDGTAKLMDPGTGRCVATLGRHTSWRSVAFAPDGL 999
QY 1100 FSTSDAKTAKIWSFDLLPLHLRGNHGCVRCSAFSDVSTLLATGDDNGEIRINWVNSG 1159
DB 1000 LASGSDGTARINDTGTGCLQTLAGHTYLTICSVAFSLDGLQGLASGSDQTLRLMEVQT 1059
QY 1160 ELLHLCLAPLEEGAATHGGVTDLCFSPDGMKMLISAGG--YIKWNVVVTGESSQTFYNG 1217
DB 1060 ACLR---TLTEK-----TGMVFLAFSPDGOILASGSDMTVKLWQVGTGRCVKTLGPH 1111
QY 1218 TNLKKIHWSPDFT 1231
DB 1112 SLVVSIAVAPDGT 1125

RESULT 11
YY46 ANASP
ID YY46 ANASP STANDARD; PRT; 1526 AA.
AC Q8YR1:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical WD-repeat protein alr3466.
GN OrderedLocusNames=alr3466.
OS Anabaena sp. (Strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
CX NCBI_TaxID=103690;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=2159285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
CC -!- SIMILARITY: Contains 1 pentapeptide repeat domain.
CC -!- SIMILARITY: Contains 16 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```


Query Match 10.2%; Score 677.5; DB 2; Length 1711;
 Best Local Similarity 22.8%; Pred. NO. 2.5e-30;
 Matches 322; Conservative 248; Mismatches 527; Indels 315; Gaps 64;

QY 7 NCLQHREALEKDKTSYIMDHMSDGLTITSEBEKVRNEPT-----QQQ 51
 DB 359 NC-LQSR-LTEHQONLATTLL---YLRKLAESPTEGPEVLVGENIQLNRWLROOL 413

QY 52 RAAMLIRKML-----KKONDSYVSFYNALHGYKDLAALLHDGIPVSSSSGKDSVSGIT 107
 DB 414 RRSLSYQIIFDCPLDKHSYIS-----LDQWVEDLQGFEGQGCIIIAAASSPENPQOFL 468

QY 108 SYRTVL-----CEGVPPQRPVVFVTRKLVNA 135
 DB 469 QTLHSTLQAOHQEPNLAAAWINQLSSPLHLWLSGAKGVIEIIPASTDAKGRPNA 528

QY 136 I-----CQKSLKLGEPGWTHGMAGKGSV--- 162
 DB 529 IVDLIGICPYRGLQAFQBEDVOYFGRETLTQOLADLANK-SFMAVVGASGGKSVVQA 587

QY 163 -LAEAVRDHSLLBGCFFGGVHVSVGKQDKSGLLMKLQN-LCTRLODQESFQRLPLN- 219
 DB 588 GLIAQLRGQQL-----PGSQEWMMKSPRPGEPYPLVSLSHCLVDSGTEKEKAYQOMLEG 642

QY 220 --IBEAKDRILMLRKHPRSLLILDVMDSWVLKAFDSQCQIILT-----TRDK--- 267
 DB 643 MLYOGAOGFVHLHRRPVMVLVDQPEELFTLAASEDRQRFDTVLGALELSPDKPKL 702

QY 268 --SVTDSVMGPKYVVPVSSIGKEGL-----ELSLFVN-MKKADL---PEQAH 312
 DB 703 IVTLRAPFACLEIPTLAKLQSSILLPCLTQOEYRRIIHPAEKVGTVDPVELVEV 762

QY 313 IIEKCKSPVVSILGALLRDPFRWEY-----YKQLQWQKQPKRKRSY 359
 DB 763 LLOELHNSPGDLPLEFVLEQL---WEYRDKGVITLQAYQYLGIGK---ALEKKAQ 815

QY 360 DYEALD-EAMS-----ISVEMLRD-----IKDYTLDSILQDKVQVPTK 398
 DB 816 VYDTLDSEAQCTRWIFLSLTQLGEGTDRRRVLKSELIVKKY--PVALVERTLQVLT 873

QY 399 VLCILMDMETEE-----VEDILQ-----EFVYKSLPCDRNGKSFYYLH 438
 DB 874 AKLVVNGEWEEGKQKQAGRGQGENILUTTSVTIEVAHEVII---RYWSTLRWL- 929

QY 439 DLQVDFLTEKNCSQLDLHKIITQFORHYQPHLTSPD-----QEDCM 481
 DB 930 -----EENRSLRS-HRQ-IEQSAALWQNNQPDFLQGVRLAEAEIYLYNDEL 979

QY 482 YW--YNFL--AYHWASAKMHKELCALMPSLDWIKAKTEL-----VGPALHIEFVYRHIL 533
 DB 980 SWDVQHFTEACLHERRRQKQEQSRQAQRAVSIISTGLTAFGLAVFAVYQQTQNAQLK 1039

QY 534 DEKDC-AVSENP-----QEFSLNGLHLCRQFPFNVOLGL---CEPETSEVYQAKLQAK 585
 DB 1040 BQALNSLNFSLHSHQLEALMTSVQAGKEVQNI-SLIGPPDTRTQTATLQOAVNTQ 1098

QY 586 QEVNDGMLYLEWINKNITNLSRLVVRPHTDAVYHACFSEDOQIASCGADKTLQVFAE 645
 DB 1099 ER--NRLHNAW-----VTSVS-----YSPGGEVIAGSDVNTIHLWRD 1136

QY 646 TGEKLBIKAHEDVLCCAFSTDTRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNC 705
 DB 1137 -GKLLTTLTGNDGVNSVSPDCEILLASASADSTIKLW-QRNGQLITLKGHDQGVKS 1194

QY 706 HFTNSSHLLLATGSSDCFLKWLNDLNOKECHTWFHTNSVNHCRFPDDKLLASCADG 765
 DB 1195 SPSFNGE--IIASGSSDHTINLWSRAGK-LLLSLNGHSGQVNSIKFSPGDTIASADG 1251

QY 766 TLKLWATSANERKSINVKQFFLNLEDPQEDMEVIVKCCSWADGATMVA-ANKKIFLF 824
 DB 1252 TIRLW---SLDGRPLITI-----PSHTKQVL---AVTFSPDQOTIVSAGDNTVKLW 1297

QY 325 DIHTSGLLGEIHTGHSTIYQCDPSPQNHVALVSLQYCVLWNTDSRSKVADCRGLSW 884
 DB 1298 S--RNGTLTLLTEGNEAVMQVIFSPDGRLIATASADKTIITLWSRDG-NILGTFAHNE 1354

QY 885 VHGWFSPDGSFSTSSDDQTIRLWETKVKCKSAVMLKQEVVVFQENE---VMVLAVD 941
 DB 1355 VNSLSFSPDGNILASGDDNTVRLWTVNRTLPKTFYGHKGSVSVYRPSNDGKKTSLSTD 1414

QY 942 HIRLQLINGRTGQIDYLTEAQVSCCLSPHLYAFDENGAIIBELVNNR----IPQS 998
 DB 1415 STMTKWSLDGKLLQTLSPPLPDVTSISFTPNKIVALASPDHTIHLV---NRQGLLRS 1470

QY 999 RFQHKVTVHIOFTADEKTLISSDDAEIQW--NWQDKCIFLRGHQETVKDFPLKNS 1056
 DB 1471 LPGHNHWTISLSPFNKQILASGADKTIKLSVNGRLKT--LIGHNGWVTDIKFSADG 1528

QY 1057 R-LLSWSPDGTGVKVNIIITGNKEKDFVCHQCTVLSCDISHDATKFSSTSAKTAIWSFD 1115
 DB 1529 KNIYASADKTIKWS-LDGRLIITLQHSASVSNVLSPDGOTLASTSQDETIKLWNL 1587

QY 1116 LLLPHLHLRHNCGVRCSAFSDVSTLLATGDNGEIRIWNYSNGELLHLCAPLSEGAAT 1175
 DB 1588 GEL-IYTLRGHSDVYVNLSPSPDGTKTIASASDDGTIKLWVNPNGTLK-----TFQG--- 1638

QY 1176 HGGWTDLCSPDCKMLISAGGY---IKWNV 1204
 DB 1639 HRGGRSVSPDCKILAS-GGHTTVKVNWL 1669

RESULT 14
 YL24 ANASP
 ID YL24 ANASP STANDARD; PRT; 1683 AA.
 AC QYV57;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hypothetical WD-repeat protein all2124;
 GN OrderedLocusNames=all2124;
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 CX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yanada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RL DNA Res. 8:205-213(2001).
 CC -1- SIMILARITY: Contains 13 WD repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AP003588; BAB73823.1; -;
 DR PIR; AF2071; AF2071.
 DR HSP; P16649; 1ERJ.
 DR InterPro; IPR004106; Peptidase S9A.N.
 DR InterPro; IPR011047; Quin_alc_DH_like.
 DR InterPro; IPR011680; WD40.
 DR Pfam; PF00400; WD40; 14.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR SMART; SM00320; WD40; 14.
 DR PROSITE; PS00678; WD_REPEATS_1; 3.
 DR PROSITE; PS00682; WD_REPEATS_2; 14.

DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Complete proteome; Hypothetical protein; Repeat; WD repeat.
 FT REPEAT 1070 1109 WD 1.
 FT REPEAT 1111 1150 WD 2.
 FT REPEAT 1153 1192 WD 3.
 FT REPEAT 1195 1236 WD 4.
 FT REPEAT 1280 1318 WD 5.
 FT REPEAT 1320 1359 WD 6.
 FT REPEAT 1361 1400 WD 7.
 FT REPEAT 1402 1441 WD 8.
 FT REPEAT 1444 1483 WD 9.
 FT REPEAT 1486 1525 WD 10.
 FT REPEAT 1528 1567 WD 11.
 FT REPEAT 1570 1609 WD 12.
 FT REPEAT 1611 1650 WD 13.
 SQ SEQUENCE 1683 AA; 185459 MW; F7CB361FF54F7137 CRC64;
 Query Match 10.2%; Score 672; DB 1; Length 1683;
 Best Local Similarity 23.2%; Pred. No. 5.1e-30;
 Matches 281; Conservative 196; Mismatches 490; Indels 244; Gaps 41;
 QY 95 SSSGKDSVSGITSVRTVLCGQVGPORPVFVRKKLVNAIQKLSKLKGEPEGWYTHG 154
 DB 553 ASGSGKSS-----VQAGLIPO-----LRQGHIPNSEQWGIKTIR--PGVNPLEA 596
 QY 155 MACGKGVLAIAVADHSLLEGCF-----PGVHWSVGKQDKSGLLM-----KLQNLCTRL 206
 DB 597 LAR-----KLGEWGTHLLIEGMLHOGVEGVVYLRSLPQGVTVLVIDQFEELFTLAPT 651
 QY 207 DODESFQRPLNIEAKRLRLILMRKHPRLSLLILDVDSWVLKAFDSQCQILLTTRD 266
 DB 652 DR-ELFLELLGAVQYAGDRFKLII-----TLRA----- 679
 QY 267 KSVTDSVMGPKVYVVESSLKEGLEILSLFVNMMKADLPQEAHIIKECK--GSPLVV 324
 DB 680 -----DFIAPCLEVPALAEALQVAVSLV--PEKLSLDYRRVLHPAQVGLQVERA 728
 QY 325 SLIGALLR-----DPNREWYVYKQLOKQFKRIRKSSVD----- 360
 DB 729 ELVEVLLRELNSVCDLP--LLEFVLEQLQWQRTAGKLTLSQVQSLGGIKGALERSCOGV 787
 QY 361 YEALDEAMS-----TSVEMLEDKD-----YVTDL-----SILQKDVKVPKVL 401
 DB 788 YESLPPOLECAKWFISLTQLEGTEDETRRIYKSDLIIVKYFAGLVETGLNVLINAKL 847
 QY 402 ILWDMETEVEDILOEFVNKSLIFCDNKGKSPRYVHDLQVDFTLNKCNCSOLOLHKKII 461
 DB 848 VVINLEAE-----IEAQGKSFSTPPNLSLTPFVVEVAHEIILIRHWSLT 891
 QY 462 TQFQRYHQHTLSPDQ--EDCWVY-----NPLAYHMASAMHKLALMFLSDWIKAK 513
 DB 892 RHWLEENRDLRKQRIINHACIQWQSGKQADFL---LQGARLAABEDIYIYWTDELGAD 948
 QY 514 T-ELVGP--AHLIHFVEYRHILDEKDCAV-----SENFOEFL 548
 DB 949 VQEFIGACLAERKIQOLQAKNRKQARAVALSVLGIASVSFGGLAYWQGREAQFRETA 1008
 QY 549 SLN-----GHLLGRQFPFNIVQLGCEPSETSEVYQAKLQAKQEVNDGMLYLEINKNIT 604
 DB 1009 ALNSSSQANLLSHQOLAALI-----ASLKAQAQVNHVIAVPPNNKLATVT 1053
 QY 605 NLSRLV-----VRPHDAVYHACFSEDCQRIASCADKTLQVFKAAETGKLLIEIKAK 656
 DB 1054 TLQOALFEMQERNLEHGDKGVISISISRDGQTIASGSLDKTIKLWSDR--GRLFRTLNHG 1112
 QY 657 EDEVLCCAFSTDDRFATCSVDKVKIWNMTGELVHYVDSHVSQVNCCHFTNSSHLLH 716
 DB 1113 EDVAVSVSFGPDQTIASGSGSDKTIKLQTSQGLTKITTCHEQTVNVVYFSPDGN--L 1170
 QY 717 ATGSSDCFLKLWLDNQECRNTMFGHTNSVNHCRSPDKLLASCDSGTGLKWDATSN 776
 DB 1171 ASASSDHSIKLWDTTSQGLLMTLTGHSAGVITVRFSPDQTIIAGSDEKTVKLWHRQDK 1230

QY 777 ERKSINVKQFFLNLEDPQEDMEVIVKCKSWSADGARINVAANKKIFLFDIHTSGLIGEIH 836
 DB 1231 LLKTLNGHQDWN-----SLSFSPDGKTLASASADKTIKLMRIADGKLVKTL 1277
 QY 837 TGHSTIOYCDFSPQNHAAVALVALSVQCVELMNTDSRSKVDRCRCHLSWVHGVNFPDSS 896
 DB 1278 KGHNSVDVNFNSDGGKATASARDNTIKLWNRHG--IELETFTHSGSGVAVVNFPLDSNI 1336
 QY 897 FLTSSDDOTRLME-----TKKVCNKAAMLMKQEVDFVQENEVMLAVDHRIRLOLIN 950
 DB 1337 TASASLNTIETLMQRPILISPLEVLAGNSGYA-----VSPFHGSIATAGADGNIQLWH 1391
 QY 951 GRTGQIDYLTEAQVSCCLLS--PHLOYAFDENGALIEILELVNRRIFQSRFQHKKTWWH 1008
 DB 1392 SDGSLKLTIPGNKAIYGISFTPDGDLIASANADKTVKIWRVDRGKALKTLIGHNEVNK 1451
 QY 1009 TOFTADEKTLISSDDABEIOVNWQDKCIF---LRHQETVKDPRLLKNRSL--SWSPD 1064
 DB 1452 VNFSPDGKTLASARDNTVKLWNVSDGK--FKKTLKGTDEVFVWFSFDPGKIIASAD 1509
 QY 1065 GTVVMNIITGNKEKDFVCHQGTVLSCDISHDATKFSSTSDAKTIW--SPDLLLPLHEL 1123
 DB 1510 KTIRLMDSFGNLIKSLPAHNDLVYSVNFNPDGSLASTSADKTVKLRSHDGHIL--LHTF 1568
 QY 1124 RHNGCVRCSAFVSVDSTLLATGDDNGEIRIWNVNGELLHLCAPLSEGAATHGQWVTDL 1183
 DB 1569 SGHNVVYSSFSFDPGRIYASASEDKTVKIWDIG---HLLTTLPO-----HQAGVMSA 1619
 QY 1184 CFSFDPGKMLIS 1194
 DB 1620 IFSPDGKTLIS 1630
 RESULT 15
 Q7ND05 PRELIMINARY; PRT; 1193 AA.
 AC Q7ND05:
 DT 01-WAR-2004 (TrEMBLrel. 26, Created)
 DT 01-WAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE WD-repeat protein.
 GN OrderedocushNames=glr2821;
 OS Gloeobacter violaceus.
 OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
 CX NCBI_TaxID=33072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCC 7421;
 RX MEDLINE=22977040; PubMed=14621292;
 RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
 RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
 RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
 RT cyanobacterium that lacks thylakoids.";
 RL DNA Res. 10:137-145(2003).
 CC -!- SIMILARITY: Contains 14 WD repeats.
 DR EMBL; AP006578; BAC90762.1; -.
 DR GO; GO:0042829; P:defense.response to pathogen; IEA.
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 14.
 DR PRINTS; PR00364; DISEASERSIST.
 DR PRINTS; PR00320; GPROTEINREPT.
 DR ProDom; PD000018; WD40; 13.
 DR PROSITE; PS00678; WD_REPEATS_1; 11.
 DR PROSITE; PS00682; WD_REPEATS_2; 14.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Complete proteome; Repeat; WD-repeat.
 SQ SEQUENCE 1193 AA; 130908 MW; 07AB8750D470E6C3 CRC64;
 Query Match 10.1%; Score 666; DB 2; Length 1193;
 Best Local Similarity 23.1%; Pred. No. 6.9e-30;



GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 18:25:54 ; Search time 176 Seconds
(without alignments)
2543.717 Million cell updates/sec

Title: US-10-646-396-2

Perfect score: 6619

Sequence: 1 MDKARNCLLQHREALEKDI.....FKTYVTVDNLGILYILQTL 1248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729239 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_23Sep04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6619	100.0	1248	4	AAy97636
2	6619	100.0	1248	5	ABg65661 Human apo
3	6619	100.0	1248	6	ABP72163 Human APA
4	6616	100.0	1248	4	AAy97646 Apaf-1XL-
5	6616	100.0	1248	4	AAy97639 Apaf-1XL-
6	6616	100.0	1248	4	AAy97643 Apaf-1XL-
7	6615	99.9	1248	4	AAy97644 Apaf-1XL-
8	6614	99.9	1248	4	AAy97645 Apaf-1XL-
9	6614	99.9	1248	4	AAy97641 Apaf-1XL-
10	6613	99.9	1248	4	AAy97647 Apaf-1XL-
11	6611	99.9	1248	4	AAy97642 Apaf-1XL-
12	6548.5	98.9	1237	4	AAy97638 Apaf-1WD
13	6548.5	98.9	1237	7	ADP65308 Human apo
14	6352.5	96.0	1205	2	AAW91072 Apoptosis
15	6352.5	96.0	1205	4	AAy97637 Apaf-1L p
16	6282	94.9	1194	2	AAW91071 Apoptosis
17	6282	94.9	1194	3	AAy49432 Human ful
18	6282	94.9	1194	4	AAy97649 Wild type
19	6282	94.9	1194	5	ABg65662 Human apo
20	6282	94.9	1194	6	ABB82741 Human Apa
21	6282	94.9	1194	7	ADP65303 Human apo
22	6282	94.9	1194	8	ADG87139 Human Apa
23	6112	92.3	1149	4	AAy97640 Apaf-1XL(
24	6003.5	90.7	1199	8	ADM87180 Human pro
25	5933.5	89.6	1526	8	ADM87179 Human pro

26	5817	87.9	1238	5	ABG65660	Abg65660 Mouse apo
27	3205.5	48.4	769	4	ABG13410	Abg13410 Novel hum
28	2977	45.0	570	4	AAy97648	AAy97648 Apaf-1XL(
29	1589	24.0	308	5	ABJ04764	ABJ04764 Apaf-1 pr
30	492	7.4	97	5	ABB81757	ABB81757 Tumour ne
31	476	7.2	93	5	ABJ04761	ABJ04761 Apaf-1 pr
32	423	6.4	2629	2	AAW55885	AAW55885 Rat telom
33	423	6.4	2629	7	ADD21418	ADD21418 Rat TERT
34	423	6.4	2629	7	ADD21417	ADD21417 Rat TERT
35	423	6.4	2629	7	ADD21424	ADD21424 Mouse TER
36	423	6.4	2629	7	ADe63119	ADe63119 Rat Prote
37	413	6.2	2627	7	ADe63121	ADe63121 Human Pro
38	410.5	6.2	2627	2	AAW61347	AAW61347 Human tel
39	410	6.2	2630	7	ADK40934	ADK40934 Novel hum
40	405	6.1	2625	2	AAW55887	AAW55887 Human tel
41	399.5	6.0	2629	2	AAW61348	AAW61348 Mouse tel
42	346	5.2	341	3	AAg38744	AAg38744 Arabidops
43	342.5	5.2	361	4	ABB68576	ABB68576 Drosophil
44	337.5	5.1	520	5	ABP73383	ABP73383 Candida a
45	337	5.1	797	5	ABP73371	ABP73371 Candida a

ALIGNMENTS

RESULT 1

AAy97636

ID AAy97636 standard; protein; 1248 AA.

AC AAy97636;

DT 20-APR-2001 (first entry)

DE Apaf-1XL protein sequence.

KW Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;

KW apoptosis signalling pathway; cancer; autoimmune disease; variant;

KW hereditary disease; Apaf-1XL.

OS Homo sapiens.

PN WO200100827-A1.

PD 04-JAN-2001.

PF 30-JUN-2000; 2000WO-US018039.

PR 30-JUN-1999; 99US-014718P.

PA (UNMI) UNIV MICHIGAN.

PI Nunez G, Hu Y;

DR WPI; 2001-112454/12.

DR N-PSDB; AAA91115.

PT Compositions for screening apoptosis pathway agonists and antagonists, useful for the treatment and diagnosis of cancer and autoimmune diseases, comprises new splice variants of wild-type Apaf-1 gene.

PS Disclosure; Fig 25; 101pp; English.

CC This sequence represents an Apaf-1 variant of the invention. The variants of the invention are all splice variants of a wild-type Apaf-1 gene (a mammalian homolog of CED-4 that participates in cytochrome c-dependent activation of caspase-3). The DNA is useful in screening assays for identifying apoptosis signalling pathway (antagonists, which are in turn useful as potential therapeutic and diagnostic or prognostic tools for diverse types of cancers, autoimmune diseases and hereditary diseases and for screening compounds that modulates the interaction of Apaf-1 with other members of the signalling pathway, i.e. their substrates or ligands. The protein encoded by the Apaf-1 variant genes are useful in a cell-free assay system and the antibody generated to the translation

CC	product are used in immunoprecipitation assays to isolate new Apaf-1									
CC	pathway constituents or their natural mutants									
XX										
SQ	Sequence 1248 AA;									
	Query Match	100.0%;	Score 6619;	DB 4;	Length 1248;					
	Best Local Similarity	100.0%;	Pred. No. 0;							
	Matches 1248;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;					
QY	1	MDAKARNCLLOHREALEKDIKTSTYMDHMSIDGFLTISEBEKVNRNEPTQOQRAAMLKMI	60							
DB	1	MDAKARNCLLOHREALEKDIKTSTYMDHMSIDGFLTISEBEKVNRNEPTQOQRAAMLKMI	60							
QY	61	LKKNDSSVSYFNALLHEGYKDLAALLHDGIPVVSSSSGKDSVSGITSYVTVLCEGGVP	120							
DB	61	LKKNDSSVSYFNALLHEGYKDLAALLHDGIPVVSSSSGKDSVSGITSYVTVLCEGGVP	120							
QY	121	QRPVFWTRKLVNAIQOKLSKLKGEQWVTIHGMACGKSVLAABAVRDSHLEGCFPG	180							
DB	121	QRPVFWTRKLVNAIQOKLSKLKGEQWVTIHGMACGKSVLAABAVRDSHLEGCFPG	180							
QY	181	GVHWSVGKQDSGLLMLKQLNCTRLDQDSFSORLPINIEEAKDRLRLMLRKHPSLL	240							
DB	181	GVHWSVGKQDSGLLMLKQLNCTRLDQDSFSORLPINIEEAKDRLRLMLRKHPSLL	240							
QY	241	ILDVDVNSVLAQFOSQOILLITRDKSVYDVSVMGPKYVVPVVESSLGKGLLEILSLFVN	300							
DB	241	ILDVDVNSVLAQFOSQOILLITRDKSVYDVSVMGPKYVVPVVESSLGKGLLEILSLFVN	300							
QY	301	MKKADLPQAHSIKECKGSLVWSLIGALLRDFPNWEYIYKQLONQKQPKRINKSSSYD	360							
DB	301	MKKADLPQAHSIKECKGSLVWSLIGALLRDFPNWEYIYKQLONQKQPKRINKSSSYD	360							
QY	361	YEALDEAMSI SVMLRREDIKDYITDLSILOKDKVPTKVLCLMDMETEEVEDILOEFVN	420							
DB	361	YEALDEAMSI SVMLRREDIKDYITDLSILOKDKVPTKVLCLMDMETEEVEDILOEFVN	420							
QY	421	KSLLFCDRNGKFRYYLHDLOVDFLETKNCSQOLDLHKIITQFORHQPHILSPDOEDC	480							
DB	421	KSLLFCDRNGKFRYYLHDLOVDFLETKNCSQOLDLHKIITQFORHQPHILSPDOEDC	480							
QY	481	MYWYNFLAYHMASAKWHKELCALMFLSDWIKATELVGPAHLIHFVEYRHLIDKCAV	540							
DB	481	MYWYNFLAYHMASAKWHKELCALMFLSDWIKATELVGPAHLIHFVEYRHLIDKCAV	540							
QY	541	SENFQEFSLNGHLGROPPFNIVQLGCEPSETSEVYQAKIQAQKQEVDMGMLYLEWINK	600							
DB	541	SENFQEFSLNGHLGROPPFNIVQLGCEPSETSEVYQAKIQAQKQEVDMGMLYLEWINK	600							
QY	601	KNITNLSRLVVRPHTDAVYHACFSEDQRIASCADKTLQVFKATGSKGLEIKAHEDEV	660							
DB	601	KNITNLSRLVVRPHTDAVYHACFSEDQRIASCADKTLQVFKATGSKGLEIKAHEDEV	660							
QY	661	LCCAFSTDDRFTATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLATGS	720							
DB	661	LCCAFSTDDRFTATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLATGS	720							
QY	721	SDCFLKLDLNQKCEKNTMGHTNSVNHCFSPDDKLLASCADGTIKLWDATSAHERKS	780							
DB	721	SDCFLKLDLNQKCEKNTMGHTNSVNHCFSPDDKLLASCADGTIKLWDATSAHERKS	780							
QY	781	INVKQFFLNLEDQEDMEVIVKCCSWADGARIMWAANKIPLFDIHTSGLLGEIHTGHH	840							
DB	781	INVKQFFLNLEDQEDMEVIVKCCSWADGARIMWAANKIPLFDIHTSGLLGEIHTGHH	840							
QY	841	STIQVCDSPONHVALVALSOYVELWNTDSRSKVDACRGLSHVHGVMSPDGSFLTS	900							
DB	841	STIQVCDSPONHVALVALSOYVELWNTDSRSKVDACRGLSHVHGVMSPDGSFLTS	900							
QY	901	SDQOTIRLWETKVKCKNSAVMLKQEVDDVVFQFQENVMVLAVDHIRRLQINGRTGQIDYLT	960							
DB	901	SDQOTIRLWETKVKCKNSAVMLKQEVDDVVFQFQENVMVLAVDHIRRLQINGRTGQIDYLT	960							

RESULT 2
 AEG5661
 ID AEG5661 standard; protein; 1248 AA.
 XX
 AC AEG5661;
 XX
 DT 26-AUG-2002 (first entry)
 XX
 Human apoptotic protease activating factor 1 (Apaf-1) exon 1 3'end.
 DE
 XX
 Antisense compound; apoptotic protease activating factor 1; Apaf-1;
 KW hyperproliferative disorder; cancer; breast cancer; colon cancer;
 KW haematopoietic cancer; prostate cancer; antisense gene therapy;
 KW infection; inflammation; tumour formation; antisense technology.
 XX
 OS Homo sapiens.
 XX
 WO200232921-A1.
 XX
 25-APR-2002.
 XX
 15-OCT-2001; 2001WO-US032116.
 XX
 16-OCT-2000; 2000US-00690364.
 XX
 (ISIS-) ISIS PHARM INC.
 XX
 Zhang H, Watt AT;
 WPI: 2002-463303/49.
 XX
 N-PSDB; ABK93570.
 XX
 Novel antisense compound that hybridizes and inhibits nucleic acid
 encoding apoptotic protease activating factor 1, for treating
 hyperproliferative disorder e.g. cancer, preferably breast, colon, or
 prostate cancer.
 XX
 Disclosure; Page 109-114; 138pp; English.
 XX
 The invention describes an antisense compound (I) 8-50 nucleobases in
 length targeted to a nucleic acid molecule (II) encoding an apoptotic
 protease activating factor 1 (Apaf-1), where (I) specifically hybridises
 with and inhibits expression of Apaf-1, or specifically hybridises with
 at least an 8-nucleobase portion of an active site on (II). (I) is useful
 for inhibiting the expression of Apaf-1 in cells or tissues, and for
 treating an animal having a disease or condition associated with Apaf-1,
 where the disease or condition is a hyperproliferative disorder such as
 cancer, preferably breast, colon, haematopoietic or prostate cancer. (I)
 is also useful for diagnostics, therapeutics, prophylaxis, as research
 reagents and kits, for distinguishing functions of various members of a
 biological pathway, and in antisense gene therapy. (I) is also useful
 prophylactically, e.g. to prevent or delay infection, inflammation or

CC tumour formation. This is the amino acid sequence of an Apaf-1 protein
 CC DNA encoding which used in the creation of antisense oligonucleotides for
 CC modulating the activity and expression of Apaf-1
 XX
 SQ Sequence 1248 AA;

Query Match	100.0%;	Score 6619;	DB 5;	Length 1248;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 1248;	Conservative 0;			Gaps 0;

Qy	1	MDAKARNCLLOHREALEKDIKTSYIMDMHISDGLTISEEEKVNRNEPTQOQRAAMLIKMI	60
Db	1	MDAKARNCLLOHREALEKDIKTSYIMDMHISDGLTISEEEKVNRNEPTQOQRAAMLIKMI	60
Qy	61	LKQNDSSVSYFNALLHGGYKDLAALLHDGIPVSSSSGKDSVSGITSYVTVLCEGGVP	120
Db	61	LKQNDSSVSYFNALLHGGYKDLAALLHDGIPVSSSSGKDSVSGITSYVTVLCEGGVP	120
Qy	121	QRPVVVTRKKLVNAIQOKLSKLGEPGWVTHGAGCGKSVLAABAVRDHSLLEGCPG	180
Db	121	QRPVVVTRKKLVNAIQOKLSKLGEPGWVTHGAGCGKSVLAABAVRDHSLLEGCPG	180
Qy	181	GVHVSVGKQKSGLLMKLQNLCTRLDQESFQSLPLNIEEAKDRLRLMLRKHPRLSLL	240
Db	181	GVHVSVGKQKSGLLMKLQNLCTRLDQESFQSLPLNIEEAKDRLRLMLRKHPRLSLL	240
Qy	241	ILDVDWDSWLKAFDSQCOILLTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN	300
Db	241	ILDVDWDSWLKAFDSQCOILLTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN	300
Qy	301	MKKADLPQAHSHIIECKGSLVSLGALLRDPNRPWEYVLKQLONKQFRRIRKSSSYD	360
Db	301	MKKADLPQAHSHIIECKGSLVSLGALLRDPNRPWEYVLKQLONKQFRRIRKSSSYD	360
Qy	361	YEALDEAMSISVEMLRDIDKYITDLSILOKDVKPTKVLCLMDMETEEVEDILOEPVN	420
Db	361	YEALDEAMSISVEMLRDIDKYITDLSILOKDVKPTKVLCLMDMETEEVEDILOEPVN	420
Qy	421	KSLLFCDRNGKSFYRLHDLQVDFLTKNCSQLODLHKKIIITQFQYHQPHTLSPDQEDC	480
Db	421	KSLLFCDRNGKSFYRLHDLQVDFLTKNCSQLODLHKKIIITQFQYHQPHTLSPDQEDC	480
Qy	481	MYWTFNFLAYHMASAKMHELCALMFSLDWIKAKTELVGPAHLIHEFVEYRHLIDKOCV	540
Db	481	MYWTFNFLAYHMASAKMHELCALMFSLDWIKAKTELVGPAHLIHEFVEYRHLIDKOCV	540
Qy	541	SENFQEFSLNGLHLLGRQFFNIVOLGLCEPETSEVYQOAKLOAKQEVNDGMLYLEWINK	600
Db	541	SENFQEFSLNGLHLLGRQFFNIVOLGLCEPETSEVYQOAKLOAKQEVNDGMLYLEWINK	600
Qy	601	KNITNLSRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFKAEKGKLEIKAHEDV	660
Db	601	KNITNLSRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFKAEKGKLEIKAHEDV	660
Qy	661	LCCAFSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS	720
Db	661	LCCAFSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS	720
Qy	721	SDCFLLKLDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS	780
Db	721	SDCFLLKLDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS	780
Qy	781	INVKOFFLNLEDPOEDMEVIVKCCSWSDAGGARIMVAANKIIFLFDIHTSGLLGEIHTGHH	840
Db	781	INVKOFFLNLEDPOEDMEVIVKCCSWSDAGGARIMVAANKIIFLFDIHTSGLLGEIHTGHH	840
Qy	841	STIQYCFSPQNHVALVALSOYVELMNTDTSRKSVDRCRGLHSWVHGMFSPDSSFLTS	900
Db	841	STIQYCFSPQNHVALVALSOYVELMNTDTSRKSVDRCRGLHSWVHGMFSPDSSFLTS	900
Qy	901	SDDOTIRLWETKVKCKNSAVMLKQEVVDVFOENEMVLAVDHIRRLQINGRTGQIDYLT	960
Db	901	SDDOTIRLWETKVKCKNSAVMLKQEVVDVFOENEMVLAVDHIRRLQINGRTGQIDYLT	960

Qy	961	EAQVSCCCLSPHLQVIAFGDENGAIIELELVNRRIFQSRPOHKKTVMHIQFTADEKTLIS	1020
Db	961	EAQVSCCCLSPHLQVIAFGDENGAIIELELVNRRIFQSRPOHKKTVMHIQFTADEKTLIS	1020
Qy	1021	SSDDAEIQVWNWQDKCIFLRGHOETVKDFRLLKNSRLLSWDFGTGVKVMNIITGNKEK	1080
Db	1021	SSDDAEIQVWNWQDKCIFLRGHOETVKDFRLLKNSRLLSWDFGTGVKVMNIITGNKEK	1080
Qy	1081	FVCHQGTVLSCDISHDATKFSSTADTKATKWSFDLLLPLHLRGHNGCVRCSAFSDVST	1140
Db	1081	FVCHQGTVLSCDISHDATKFSSTADTKATKWSFDLLLPLHLRGHNGCVRCSAFSDVST	1140
Qy	1141	LLATGDDNGEIRIWNVNGELHLHLCAPLSEGAATHGGWVTDLCPSPDGKMLISAGGYIK	1200
Db	1141	LLATGDDNGEIRIWNVNGELHLHLCAPLSEGAATHGGWVTDLCPSPDGKMLISAGGYIK	1200
Qy	1201	WNVVTGESSOTFTYNTGNTLNKKIHVSPDKTYVTVDNLGILYIQTLE	1248
Db	1201	WNVVTGESSOTFTYNTGNTLNKKIHVSPDKTYVTVDNLGILYIQTLE	1248

RESULT 3

ID	ABP72163	standard; protein; 1248 AA.
XX	ABP72163;	
XX	AC	
DT	22-APR-2003	(first entry)
XX	Human APAF1, inducer of cell death.	
DE	Human; APAF1; cell death; apoptosis; neurodegenerative disease;	
KW	heart disease; cardiomyopathy; cardiant; neuroprotective; gene therapy;	
KW	gene; ss.	
XX	Homo sapiens.	
OS	WO2003004606-A2.	
XX	16-JAN-2003.	
XX	03-JUL-2002; 2002WO-US021002.	
XX	03-JUL-2001; 2001US-00898158.	
XX	(UYCO) UNIV COLUMBIA NEW YORK.	
XX	Troy CM, Shelanski ML;	
XX	WPI; 2003-210351/20.	
XX	N-PSDB; ABZ58107.	
PT	New nucleic acid encoding an inhibitor-of-apoptosis protein, useful for	
PT	treating cancer, neurodegenerative disorder or cardiomyopathy.	
PS	Disclosure; Fig 21A; 124pp; English.	
XX	The present sequence is the protein sequence for human APAF1, a protein	
CC	that induces cell death. The invention provides a nucleic acid, such as	
CC	an antisense oligonucleotide, which specifically hybridises to a nucleic	
CC	acid encoding a protein that induces cell death, especially APAF1, RAIDD	
CC	or Diablo/SMAC. A claimed method for inhibiting a cell's death	
CC	(especially a neuronal cell's death) comprises contacting the cell with	
CC	the nucleic acid under conditions permitting the nucleic acid to enter	
CC	the cell, especially the use of a vector, liposome, or a mechanical or	
CC	electrical means. The method is used to treat a neurodegenerative	
CC	disorder, especially a brain disorder or central nervous system disorder,	
CC	or a heart disorder, especially cardiomyopathy, in a human (all claimed)	
XX	Sequence 1248 AA;	

Query Match 100.0%; Score 6619; DB 6; Length 1248;

		Best Local Similarity 100.0%; Pred. No. 0;				Matches 1248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MDAKARNCLLQHREALEKDIKTSYIMDMHISDGLFATISEEKKVRNEPTQOQRAAMLKMI	60						
Db	1	MDAKARNCLLQHREALEKDIKTSYIMDMHISDGLFATISEEKKVRNEPTQOQRAAMLKMI	60						
Qy	61	LKXNDSDYVSFYFNALLHGGYKDLAALLHDLGIPVVSSSSGKDSVSGITSYVTVLCEGVP	120						
Db	61	LKXNDSDYVSFYFNALLHGGYKDLAALLHDLGIPVVSSSSGKDSVSGITSYVTVLCEGVP	120						
Qy	121	QRPVVFTRKLVNAIQKLSKLKGEFGWYTHGMAGCKSVLAARAVRHSLEGGCPG	180						
Db	121	QRPVVFTRKLVNAIQKLSKLKGEFGWYTHGMAGCKSVLAARAVRHSLEGGCPG	180						
Qy	181	GVHWSVGKDGSLMLKMLCNLTLDODESQRPLNIEAKDRILMLRKHSLL	240						
Db	181	GVHWSVGKDGSLMLKMLCNLTLDODESQRPLNIEAKDRILMLRKHSLL	240						
Qy	241	ILDDVNDSVLKAFOQCOIILLTRDKSVTDSVNGPKYVVPVSSLGKGELEILSLFVN	300						
Db	241	ILDDVNDSVLKAFOQCOIILLTRDKSVTDSVNGPKYVVPVSSLGKGELEILSLFVN	300						
Qy	301	MXKADLPEQAHGIIKECKGSLVLSLIGALLRPPNRMWYLLKQLQNKQPKIRKSSVD	360						
Db	301	MXKADLPEQAHGIIKECKGSLVLSLIGALLRPPNRMWYLLKQLQNKQPKIRKSSVD	360						
Qy	361	YEALDEAMSISVEMLRDIDKYITDLSILOKQVKVPTKVLCLMDMETEEVEDILOEFVN	420						
Db	361	YEALDEAMSISVEMLRDIDKYITDLSILOKQVKVPTKVLCLMDMETEEVEDILOEFVN	420						
Qy	421	KSLFCDNRNGKFRYYLHDLQVDFLTKNCSQQLHKKIITQFORVHQPHLSPQEDC	480						
Db	421	KSLFCDNRNGKFRYYLHDLQVDFLTKNCSQQLHKKIITQFORVHQPHLSPQEDC	480						
Qy	481	MYWYFLAYHMASAKWHELCALMFLSDWIKATELVGPAHLIHEFVEYRHLDEKCAV	540						
Db	481	MYWYFLAYHMASAKWHELCALMFLSDWIKATELVGPAHLIHEFVEYRHLDEKCAV	540						
Qy	541	SENFQEFSLNGLHILGROPPENIVQLGCEPTESEVYQAKIQAKQEVNDGMLYLEWINK	600						
Db	541	SENFQEFSLNGLHILGROPPENIVQLGCEPTESEVYQAKIQAKQEVNDGMLYLEWINK	600						
Qy	601	KNITNLSRLVVRPHTDAVYHACSEDOQRIASCGADKTLQVFKAEKLEIKAHEDV	660						
Db	601	KNITNLSRLVVRPHTDAVYHACSEDOQRIASCGADKTLQVFKAEKLEIKAHEDV	660						
Qy	661	LCCAFSTDRFATCSVDKKIWNSTGELVHTYDEHSEQVNCCHFTNSHLLATGS	720						
Db	661	LCCAFSTDRFATCSVDKKIWNSTGELVHTYDEHSEQVNCCHFTNSHLLATGS	720						
Qy	721	SDCFLKMLDLQKCRNMTFGHTNSVNHCRFPSPDDKLLASCADGTLKMDATSANERKS	780						
Db	721	SDCFLKMLDLQKCRNMTFGHTNSVNHCRFPSPDDKLLASCADGTLKMDATSANERKS	780						
Qy	781	INVKQFFNLNLEDPQEDMEVIVKCCSADGARIMVAANKKIFLFDIHTSGLLGEIHTGHH	840						
Db	781	INVKQFFNLNLEDPQEDMEVIVKCCSADGARIMVAANKKIFLFDIHTSGLLGEIHTGHH	840						
Qy	841	STIQYCDSPQNHAVVALSOYCVELWNITDSRSKADCRGHLNVHGVMSFPGSSFLTS	900						
Db	841	STIQYCDSPQNHAVVALSOYCVELWNITDSRSKADCRGHLNVHGVMSFPGSSFLTS	900						
Qy	901	SDQITRLWETKVKCNKSAVMLKQEVVDVVFQNEVWVLAVDHIRRLQINRGRTQIDYLT	960						
Db	901	SDQITRLWETKVKCNKSAVMLKQEVVDVVFQNEVWVLAVDHIRRLQINRGRTQIDYLT	960						
Qy	961	EAQVSCCLSLPHLOVIAFGDENGAIIELELVNRRIFQSRFOHKKTVWHIOTFADEKTLIS	1020						
Db	961	EAQVSCCLSLPHLOVIAFGDENGAIIELELVNRRIFQSRFOHKKTVWHIOTFADEKTLIS	1020						
Qy	1021	SSDDAEIQVNNWQDKCIFLRGHQETVKDFRLKKNRLLSSPDKTVKVNIIITGNKEKD	1080						

Db	1021	SSDDAEIQVNNWQDKCIFLRGHQETVKDFRLKKNRLLSSWSPDKTVKVNIIITGNKEKD	1080																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				</
----	------	---	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

||||| 1 MDKARNCLLQHREALEKDIKTSYIMDMHISDGLFTTISEEKKVRNEPTQOQRAAMLIKMI 60
61 LKNDSDSVSYFNALLHGGYKDLAALLHDGIPVSSSSGKDSVSGITSYVTVLCEGGVP 120
61 LKNDSDSVSYFNALLHGGYKDLAALLHDGIPVSSSSGKDSVSGITSYVTVLCEGGVP 120
121 QRPVVVTRKKLVNAIOQKLSKLKEPGWVTHGMAGCGKSVLAABAVRDSHSLLEGCPGP 180
121 QRPVVVTRKKLVNAIOQKLSKLKEPGWVTHGMAGCGKSVLAABAVRDSHSLLEGCPGP 180
181 GVHVSVCKQKSGLLMKLQNLCTELDQDESFSORLPLNIEAKDRILMLRKHPRSL 240
181 GVHVSVCKQKSGLLMKLQNLCTELDQDESFSORLPLNIEAKDRILMLRKHPRSL 240
241 ILDDVWDSWLKAFDSQOQIILLTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFVN 300
241 ILDDVWDSWLKAFDSQOQIILLTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFVN 300
301 MKKADLPQAHSIIEKCGSPVLSLIGALLRDPFNRWEYVYLKQNLQKQFRIKSSSYD 360
301 MKKADLPQAHSIIEKCGSPVLSLIGALLRDPFNRWEYVYLKQNLQKQFRIKSSSYD 360
361 YEALDEAMSISVEMLRDIDKYYTDLSTLOKDVKPTKVLCLMDMETEEVEDILOEFVN 420
361 YEALDEAMSISVEMLRDIDKYYTDLSTLOKDVKPTKVLCLMDMETEEVEDILOEFVN 420
421 KSLFLCDNRNGKSFYRLHDLQVDFTEKNCQLODLHKKIITQFORYHOPHTLSPDQDC 480
421 KSLFLCDNRNGKSFYRLHDLQVDFTEKNCQLODLHKKIITQFORYHOPHTLSPDQDC 480
481 MYWNFLAYHMASAKMKELCALMFLSDWIKAKTELUGPAHLIHEFVEYRHLDEKCAV 540
481 MYWNFLAYHMASAKMKELCALMFLSDWIKAKTELUGPAHLIHEFVEYRHLDEKCAV 540
541 SENQEFSLNGLHILGROPPFNIYVQLGLCEPETSEVYQOAKLOAQEVNDGMLYLEWINK 600
541 SENQEFSLNGLHILGROPPFNIYVQLGLCEPETSEVYQOAKLOAQEVNDGMLYLEWINK 600
601 KNIITNLSRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFKAETGKLEIKAHEDV 660
601 KNIITNLSRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFKAETGKLEIKAHEDV 660
661 LCCAFSTDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
661 LCCAFSTDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
721 SDCFLKLDLWLNQKRCNTMFGHTSVNHCPSDDKLLASCADGTLKLDATSANERKS 780
721 SDCFLKLDLWLNQKRCNTMFGHTSVNHCPSDDKLLASCADGTLKLDATSANERKS 780
781 INVQKFFLNLEDPQEDMEVIVKCCSWSDAGARIMVAANKIPLFDIHTSGLLGEIHTGHH 840
781 INVQKFFLNLEDPQEDMEVIVKCCSWSDAGARIMVAANKIPLFDIHTSGLLGEIHTGHH 840
841 STIQCYDFSQNHILAVALSQYVELMNTDSRSKVADCRGHLVHGVWVSPDSSFLTS 900
841 STIQCYDFSQNHILAVALSQYVELMNTDSRSKVADCRGHLVHGVWVSPDSSFLTS 900
901 SDDQIRLWETKVKCKNSAVMLKQEVVDVVFQENEVMLAVDHIRRLQINGRTQIDYLT 960
901 SDDQIRLWETKVKCKNSAVMLKQEVVDVVFQENEVMLAVDHIRRLQINGRTQIDYLT 960
961 EAOVSCCLSPHLOVIAFGDENGALIEILEVNNRIFQSRFOHKKTVMHIQFTADEKTLIS 1020
961 EAOVSCCLSPHLOVIAFGDENGALIEILEVNNRIFQSRFOHKKTVMHIQFTADEKTLIS 1020
1021 SSDDAEIOVNNWQDKICIFLAGHQBTVKDFRLLKNSRLLSVDFGTVKVMNIIITGNKEK 1080
1021 SSDDAEIOVNNWQDKICIFLAGHQBTVKDFRLLKNSRLLSVDFGTVKVMNIIITGNKEK 1080
1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLLPLHLRHGNCVRCSAFSDVST 1140
|||||

Db 1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLLPLHLRHGNCVRCSAFSDVST 1140
Qy 1141 LLATGDDNGEIRIWNVNGELLHLHCAPLSBEGAATHGCGWTDLCFSPDGKMLISAGGYIK 1200
Db 1141 LLATGDDNGEIRIWNVNGELLHLHCAPLSBEGAATHGCGWTDLCFSPDGKMLISAGGYIK 1200
Qy 1201 WNNVVTGESSQTFYNTGNTNKKIHSVDPFKTYVTVVDNLGILYILQTL 1248
Db 1201 WNNVVTGESSQTFYNTGNTNKKIHSVDPFKTYVTVVDNLGILYILQTL 1248

RESULT 5

AA97639
ID AAY97639 standard; protein; 1248 AA.
XX AAY97639;
XX 20-APR-2001 (first entry)
XX Apaf-1XL-M368L protein sequence.
DE Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
KW hereditary disease; Apaf-1XL-M368L.
XX Homo sapiens.
XX WO200100827-A1.
XX 04-JAN-2001.
XX 30-JUN-2000; 2000WO-US018039.
XX 30-JUN-1999; 99US-0141718P.
XX (UNMI) UNIV MICHIGAN.
XX Nunez G, Hu Y;
XX WPI; 2001-112454/12.
XX N-PSDB; AAA91118.
XX Compositions for screening apoptosis pathway agonists and antagonists,
XX useful for the treatment and diagnosis of cancer and autoimmune diseases,
XX comprises new splice variants of wild-type Apaf-1 gene.
XX Disclosure; Fig 28; 101pp; English.
XX This sequence represents an Apaf-1 variant of the invention. The variants
XX of the invention are all splice variants of a wild-type Apaf-1 gene (a
XX mammalian homolog of CED-4 that participates in cytochrome c-dependent
XX activation of caspase-3). The DNA is useful in screening assays for
XX identifying apoptosis signalling pathway (antagonists, which are in turn
XX useful as potential therapeutics and diagnostic or prognostic tools for
XX diverse types of cancers, autoimmune diseases and hereditary diseases and
XX for screening compounds that modulates the interaction of Apaf-1 with
XX other members of the signalling pathway, i.e. their substrates or
XX ligands. The protein encoded by the Apaf-1 variant genes are useful in a
XX cell-free assay system and the antibody generated to the translation
XX product are used in immunoprecipitation assays to isolate new Apaf-1
XX pathway constituents or their natural mutants
XX Sequence 1248 AA;

Query Match 100.0%; Score 6616; DB 4; Length 1248;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1247; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDKARNCLLQHREALEKDIKTSYIMDMHISDGLFTTISEEKKVRNEPTQOQRAAMLIKMI 60
Db 1 MDKARNCLLQHREALEKDIKTSYIMDMHISDGLFTTISEEKKVRNEPTQOQRAAMLIKMI 60
Qy 61 LKNDSDSVSYFNALLHGGYKDLAALLHDGIPVSSSSGKDSVSGITSYVTVLCEGGVP 120

Db 61 LKNDSDSYVFNALLHGGYKDLAALLHGGIPVVSSSGKDSVSGITSYVTVLCEGVP 120
 Qy 121 QRPVVFTRKLVNAIQOKSLKAGEPGWWTIHGMAGCGKSVLAEEAVRDSHLSLEGCFPG 180
 Db 121 QRPVVFTRKLVNAIQOKSLKAGEPGWWTIHGMAGCGKSVLAEEAVRDSHLSLEGCFPG 180
 Qy 181 GVHWVSGKODKSGLLMLKQLNCLTRLDQDESFSORPLNIIEAKDRLRLMLRKHPSLL 240
 Db 181 GVHWVSGKODKSGLLMLKQLNCLTRLDQDESFSORPLNIIEAKDRLRLMLRKHPSLL 240
 Qy 241 ILDDVDSVLAQPSQCOQILLITRDSKSVTDSVMGPKYVVPVSSLGKEKGLTSLFVN 300
 Db 241 ILDDVDSVLAQPSQCOQILLITRDSKSVTDSVMGPKYVVPVSSLGKEKGLTSLFVN 300
 Qy 301 MKKADLPKQAHSIIEKCKGSLVSLIGALLRPFNRWEYIKOLONKQFKRIRKSSYD 360
 Db 301 MKKADLPKQAHSIIEKCKGSLVSLIGALLRPFNRWEYIKOLONKQFKRIRKSSYD 360
 Qy 361 YEALDPAASVEMLRBEDIKDYITDLSILOKVVPKVKLCILMDMETEEVEDILOEFVN 420
 Db 361 YEALDPAASVEMLRBEDIKDYITDLSILOKVVPKVKLCILMDMETEEVEDILOEFVN 420
 Qy 421 KSLLCORNGKFRYYLHDLQVDFLTKNCSQLODLHKKIITQFORVHQPHILSPQEDC 480
 Db 421 KSLLCORNGKFRYYLHDLQVDFLTKNCSQLODLHKKIITQFORVHQPHILSPQEDC 480
 Qy 481 MYWYNFLAYHMASAKMHKELCALMFLSDMTKAKTELVGPAHLIHEFVEYRHLDEKCAV 540
 Db 481 MYWYNFLAYHMASAKMHKELCALMFLSDMTKAKTELVGPAHLIHEFVEYRHLDEKCAV 540
 Qy 541 SENFQFSLNGLHILGROPPFNIVQLGLCEPSETSEVYQAKLOAKQEVNDGMLYLEWINK 600
 Db 541 SENFQFSLNGLHILGROPPFNIVQLGLCEPSETSEVYQAKLOAKQEVNDGMLYLEWINK 600
 Qy 601 KNITNLRLVVRPHTDAVYHACSESDQRIASCADKTLQVFKAEKGKLEIKAHEDV 660
 Db 601 KNITNLRLVVRPHTDAVYHACSESDQRIASCADKTLQVFKAEKGKLEIKAHEDV 660
 Qy 661 LCCAFSTDDRFATICSVDKVKIWNSTGELVHTYDEHSEQVNCCHFTNSHLLATGS 720
 Db 661 LCCAFSTDDRFATICSVDKVKIWNSTGELVHTYDEHSEQVNCCHFTNSHLLATGS 720
 Qy 721 SDCLFLKLDLQKCRNTMFGHTNSVNHCFSPDDKLLASCADGTLKLDATSANERKS 780
 Db 721 SDCLFLKLDLQKCRNTMFGHTNSVNHCFSPDDKLLASCADGTLKLDATSANERKS 780
 Qy 781 INVKQFFNLDPQEDMEVIVKCCSWADGARIWAAKNKIPLFDIHTSGLLGEIHTGHH 840
 Db 781 INVKQFFNLDPQEDMEVIVKCCSWADGARIWAAKNKIPLFDIHTSGLLGEIHTGHH 840
 Qy 841 STIQYCDSPQNHVALSVOYVELWNTDSKVDACRGLSHVGMFSPDGSFLTS 900
 Db 841 STIQYCDSPQNHVALSVOYVELWNTDSKVDACRGLSHVGMFSPDGSFLTS 900
 Qy 901 SDQTIWLWETKVKCKSAVLMKQEVVDVFOENVMVLAVDHIIRLQLINGRTGQIDYLT 960
 Db 901 SDQTIWLWETKVKCKSAVLMKQEVVDVFOENVMVLAVDHIIRLQLINGRTGQIDYLT 960
 Qy 961 EAQVSCCLSPHLYQVAFDNGAIBILELVNRRIFQSRFQHKKTVMWHIOFTADEKTLIS 1020
 Db 961 EAQVSCCLSPHLYQVAFDNGAIBILELVNRRIFQSRFQHKKTVMWHIOFTADEKTLIS 1020
 Qy 1021 SSDDAEIOVMNQWLDKICIFLRGQETVKDFRLLKNSRLLSWSPDGTVKVNNITGNKEK 1080
 Db 1021 SSDDAEIOVMNQWLDKICIFLRGQETVKDFRLLKNSRLLSWSPDGTVKVNNITGNKEK 1080
 Qy 1081 FVCHQGTVLSDISHDATKFSSTADKTAKIWSFOLLPLHELGHNGCVRCFAFSDVT 1140
 Db 1081 FVCHQGTVLSDISHDATKFSSTADKTAKIWSFOLLPLHELGHNGCVRCFAFSDVT 1140
 Qy 1141 LLATGDGNGEIRIWNVNGELLHLCLPSEGAATHGCGWTDLCFSPDGKMLISAGGYIK 1200

Db 1141 LLATGDGNGEIRIWNVNGELLHLCLPSEGAATHGCGWTDLCFSPDGKMLISAGGYIK 1200
 Qy 1201 WNVVVTGSSQTFYTNCTNLKKIHVSPDFKTYVTVNDNLGILYILQTL 1248
 Db 1201 WNVVVTGSSQTFYTNCTNLKKIHVSPDFKTYVTVNDNLGILYILQTL 1248
 RESULT 6
 AAY97643
 ID AAY97643 standard; protein; 1248 AA.
 AC AAY97643;
 XX 20-APR-2001 (first entry)
 DE Apaf-1XL-E39Q protein sequence.
 KW Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 XX hereditary disease; Apaf-1XL-E39Q.
 OS Homo sapiens.
 XX WO200100827-A1.
 XX PN
 XX 04-JAN-2001.
 XX 30-JUN-2000; 2000WO-US018039.
 XX 30-JUN-1999; 99US-0141718P.
 XX (UNMI) UNIV MICHIGAN.
 XX Nunez G, Hu Y;
 XX WPI: 2001-112454/12.
 XX N-PSDB; AAA91122.
 PT Compositions for screening apoptosis pathway agonists and antagonists.
 PT useful for the treatment and diagnosis of cancer and autoimmune diseases,
 PT comprises new splice variants of wild-type Apaf-1 gene.
 XX Disclosure; Fig 32; 10pp; English.
 CC This sequence represents an Apaf-1 variant of the invention. The variants
 CC of the invention are all splice variants of a wild-type Apaf-1 gene (a
 CC mammalian homolog of CED-4 that participates in cytochrome c-dependent
 CC activation of caspase-3). The DNA is useful in screening assays for
 CC identifying apoptosis signalling pathway (antagonists, which are in turn
 CC useful as potential therapeutics and diagnostic or prognostic tools for
 CC diverse types of cancers, autoimmune diseases and hereditary diseases and
 CC for screening compounds that modulates the interaction of Apaf-1 with
 CC other members of the signalling pathway, i.e. their substrates or
 CC ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 CC cell-free assay system and the antibody generated to the translation
 CC product are used in immunoprecipitation assays to isolate new Apaf-1
 CC pathway constituents or their natural mutants
 XX Sequence 1248 AA;
 SQ
 Query Match 100.0%; Score 6616; DB 4; Length 1248;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1247; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MDKARNCLLQHREALEKDKITSYIMDMHISDGLPTISEEKVNEPTQOQRAAMLIRKI 60
 Db 1 MDKARNCLLQHREALEKDKITSYIMDMHISDGLPTISEEKVNEPTQOQRAAMLIRKI 60
 Qy 61 LKNDSDSYVFNALLHGGYKDLAALLHGGIPVVSSSGKDSVSGITSYVTVLCEGVP 120
 Db 61 LKNDSDSYVFNALLHGGYKDLAALLHGGIPVVSSSGKDSVSGITSYVTVLCEGVP 120
 Qy 121 QRPVVFTRKLVNAIQOKSLKAGEPGWWTIHGMAGCGKSVLAEEAVRDSHLSLEGCFPG 180

Db 121 QRPVVVTRKKLVNAIQKLSKLGEPGWVTHGAGCGKSVLAEEAVRDSHSLLEGCPG 180
 Qy 181 GVHWSVGKQKSGLLMKLQNLCTRLDQDQSFQSLPLNIEBEAKDRILMLRKHPSLL 240
 Db 181 GVHWSVGKQKSGLLMKLQNLCTRLDQDQSFQSLPLNIEBEAKDRILMLRKHPSLL 240
 Qy 241 ILDDVWDSWVLKAFDSQOQIILTRDKSVTSVMGPKYVVPVSSSLGKEGLEILSLFN 300
 Db 241 ILDDVWDSWVLKAFDSQOQIILTRDKSVTSVMGPKYVVPVSSSLGKEGLEILSLFN 300
 Qy 301 MKKADLPQAHSIKECKGSLPVLVSLGALLARDPENRWEYVYLKOLQKPKRIRKSSVD 360
 Db 301 MKKADLPQAHSIKECKGSLPVLVSLGALLARDPENRWEYVYLKOLQKPKRIRKSSVD 360
 Qy 361 YEALDEAMSISEVLEMDREDIKDYTTDLSTLOKDVKVPKVLGILMDMETEVEDILQEPVN 420
 Db 361 YEALDEAMSISEVLEMDREDIKDYTTDLSTLOKDVKVPKVLGILMDMETEVEDILQEPVN 420
 Qy 421 KSLFLCDRNGKSFYRLHDLQVDFLTKNCSQLODLHKKIITQFORHQPHTLSPDQEDC 480
 Db 421 KSLFLCDRNGKSFYRLHDLQVDFLTKNCSQLODLHKKIITQFORHQPHTLSPDQEDC 480
 Qy 481 MYTNFLAYHMASAKWHELCALMFSLDWIRAKTELVGPAHLIHEFVEYRHLDEKCAV 540
 Db 481 MYTNFLAYHMASAKWHELCALMFSLDWIRAKTELVGPAHLIHEFVEYRHLDEKCAV 540
 Qy 541 SENFQEFSLNGLHGRFPFNIVOLGICEPETSEVYQOAKLOKQEVNDGMLYLEWINK 600
 Db 541 SENFQEFSLNGLHGRFPFNIVOLGICEPETSEVYQOAKLOKQEVNDGMLYLEWINK 600
 Qy 601 KNIITNLSRLVVRPHTDAVYHACFSEDGQRIASCGADKTLQVFKAPTGEKLEIKAHEDV 660
 Db 601 KNIITNLSRLVVRPHTDAVYHACFSEDGQRIASCGADKTLQVFKAPTGEKLEIKAHEDV 660
 Qy 661 LCCAFSTDDRTATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGS 720
 Db 661 LCCAFSTDDRTATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGS 720
 Qy 721 SDCFLKLDLWLNQECRNTMFGHTSNVHCRSPDDKLLASCADGTLKLDATSANERKS 780
 Db 721 SDCFLKLDLWLNQECRNTMFGHTSNVHCRSPDDKLLASCADGTLKLDATSANERKS 780
 Qy 781 INVQKFFLNLEDPOEDMEVIVKCCSWSDAGARIMVAANKIIFLDIHTSGLLGEIHTGHH 840
 Db 781 INVQKFFLNLEDPOEDMEVIVKCCSWSDAGARIMVAANKIIFLDIHTSGLLGEIHTGHH 840
 Qy 841 STIQYCDSPQNLAVVALSOYCVELWNTDSRSKVADCRGHLVWGHVWVSPDGSSFLT 900
 Db 841 STIQYCDSPQNLAVVALSOYCVELWNTDSRSKVADCRGHLVWGHVWVSPDGSSFLT 900
 Qy 901 SDDQTLRLWETKVKCKNSAVMLKQEDVVQFQNEVWVLAVDHIRRLQINGRTQIDYLT 960
 Db 901 SDDQTLRLWETKVKCKNSAVMLKQEDVVQFQNEVWVLAVDHIRRLQINGRTQIDYLT 960
 Qy 961 EAQVSCCLSPHLOVIAFGDENGAEIIELELVNRRIFQSRFQHKTVWHIQFTADEKTLIS 1020
 Db 961 EAQVSCCLSPHLOVIAFGDENGAEIIELELVNRRIFQSRFQHKTVWHIQFTADEKTLIS 1020
 Qy 1021 SSDAEIQVWVWOLDKCIIFLRGHQETVKDFLLKNSRLLSWSFDGTVKVMNIIITGNKEK 1080
 Db 1021 SSDAEIQVWVWOLDKCIIFLRGHQETVKDFLLKNSRLLSWSFDGTVKVMNIIITGNKEK 1080
 Qy 1081 FVCHQGTVLSCDIHDAKFSTSDAKTAKTWSFDLLPLHELGRHNCVRCASFVDS 1140
 Db 1081 FVCHQGTVLSCDIHDAKFSTSDAKTAKTWSFDLLPLHELGRHNCVRCASFVDS 1140
 Qy 1141 LLATGDDNGEIRIWNVSGNELLHLCAPISEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1200
 Db 1141 LLATGDDNGEIRIWNVSGNELLHLCAPISEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1200
 Qy 1201 WNVVVTGSSQTFYTNNGNLKKIHVSPDPKTYVTVDNLGILYILOTL 1248

Db 1201 WNVVVTGSSQTFYTNNGNLKKIHVSPDPKTYVTVDNLGILYILOTL 1248
 RESULT 7
 AAY97644
 ID AAY97644 standard; protein; 1248 AA.
 XX AAY97644;
 AC AAY97644;
 DT 20-APR-2001 (first entry)
 XX
 DE Apaf-1XL-K63Q protein sequence.
 XX
 KW Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 KW hereditary disease; Apaf-1XL-K63Q.
 XX
 OS Homo sapiens.
 XX
 FN WO200100827-A1.
 XX
 PD 04-JAN-2001.
 XX
 PP 30-JUN-2000; 2000WO-US018039.
 XX
 PR 30-JUN-1999; 99US-0141718P.
 XX
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Nunez G, Hu Y;
 XX
 DR WPI; 2001-112454/12.
 XX
 DR N-PSDB; AAA91123.
 XX
 PT Compositions for screening apoptosis pathway agonists and antagonists,
 PT useful for the treatment and diagnosis of cancer and autoimmune diseases,
 PT comprises new splice variants of wild-type Apaf-1 gene.
 XX
 PS Disclosure; Fig 33; 101pp; English.
 XX
 CC This sequence represents an Apaf-1 variant of the invention. The variants
 CC of the invention are all splice variants of a wild-type Apaf-1 gene (a
 CC mammalian homolog of CED-4 that participates in cytochrome c-dependent
 CC activation of caspase-3). The DNA is useful in screening assays for
 CC identifying apoptosis signalling pathway (antagonists, which are in turn
 CC useful as potential therapeutics and diagnostic or prognostic tools for
 CC diverse types of cancers, autoimmune diseases and hereditary diseases and
 CC for screening compounds that modulates the interaction of Apaf-1 with
 CC other members of the signalling pathway, i.e. their substrates or
 CC ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 CC cell-free assay system and the antibody generated to the translation
 CC product are used in immunoprecipitation assays to isolate new Apaf-1
 CC pathway constituents or their natural mutants
 XX
 SQ Sequence 1248 AA;
 Query Match 99.9%; Score 6615; DB 4; Length 1248;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1247; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MDKARNCCLQHRALREKDKITSYIMDHMSDGLTISEEKVNEPTQQQRAAMLIKI 60
 Db 1 MDKARNCCLQHRALREKDKITSYIMDHMSDGLTISEEKVNEPTQQQRAAMLIKI 60
 Qy 61 LKQNDQSVSPYNALLHEGKDLAALLHDGTPVVSSSGKDSVSGITSYVTVLCEGVP 120
 Db 61 LKQNDQSVSPYNALLHEGKDLAALLHDGTPVVSSSGKDSVSGITSYVTVLCEGVP 120
 Qy 121 QRPVVVTRKKLVNAIQKLSKLGEPGWVTHGAGCGKSVLAEEAVRDSHSLLEGCPG 180
 Db 121 QRPVVVTRKKLVNAIQKLSKLGEPGWVTHGAGCGKSVLAEEAVRDSHSLLEGCPG 180
 Qy 181 GVHWSVGKQKSGLLMKLQNLCTRLDQDQSFQSLPLNIEBEAKDRILMLRKHPSLL 240

Db 181 GVHWVSQKQSGGLMKLQNLCTRLDQDSFQRLPNTIEAKDRULMLRKHPSLL 240
 Qy 241 ILDDVWDSWVLKAFDQSQCIILLTRDKSVTDSVMGPKYVVPVSSLGKGLKLELSL 300
 Db 241 ILDDVWDSWVLKAFDQSQCIILLTRDKSVTDSVMGPKYVVPVSSLGKGLKLELSL 300
 Qy 301 MKKADLPEQAHSIIKECKGSPVLSLIGALLRDPFNRMEYLYKQIQNKQKFRIRKSSVD 360
 Db 301 MKKADLPEQAHSIIKECKGSPVLSLIGALLRDPFNRMEYLYKQIQNKQKFRIRKSSVD 360
 Qy 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKQVVKVPTKVLCTILMDMETEEVEDILOEFVN 420
 Db 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKQVVKVPTKVLCTILMDMETEEVEDILOEFVN 420
 Qy 421 KSLFLCDNRNGKFRYLLHDLQVDFLETKNCSQQLDHLKKITQFORVHOPHTLSPOEDC 480
 Db 421 KSLFLCDNRNGKFRYLLHDLQVDFLETKNCSQQLDHLKKITQFORVHOPHTLSPOEDC 480
 Qy 481 MYWYFLAYHMASAKHKLALMFSLDWKATKTELAVGPAHLIHEFVEYRHLDEKCAV 540
 Db 481 MYWYFLAYHMASAKHKLALMFSLDWKATKTELAVGPAHLIHEFVEYRHLDEKCAV 540
 Qy 541 SENFQFLSLNGHLIGROPENIVOLGCEPTESEVYQAKLOAKQEVDMGMLYLEWINK 600
 Db 541 SENFQFLSLNGHLIGROPENIVOLGCEPTESEVYQAKLOAKQEVDMGMLYLEWINK 600
 Qy 601 KNTNLRLVVRPHDVAHYHACFSEDGQRIASCGADKTLQVFKAEETGEKLEIKAHEDV 660
 Db 601 KNTNLRLVVRPHDVAHYHACFSEDGQRIASCGADKTLQVFKAEETGEKLEIKAHEDV 660
 Qy 661 LCAFSTDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLATGS 720
 Db 661 LCAFSTDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLATGS 720
 Qy 721 SDFCLKLMDLQKCRNTMFGHTNSVNHCRFPSPDDKLLASCSADGTLKMDATSANERKS 780
 Db 721 SDFCLKLMDLQKCRNTMFGHTNSVNHCRFPSPDDKLLASCSADGTLKMDATSANERKS 780
 Qy 781 INVKOFFNLNEDPQEDMEVIVKCCSWSADGARIMAAKNKIFLFDIHTSGLLGEIHTGHH 840
 Db 781 INVKOFFNLNEDPQEDMEVIVKCCSWSADGARIMAAKNKIFLFDIHTSGLLGEIHTGHH 840
 Qy 841 STIQYCDFSPQNLAVVALSQYCVELWNTDSRSKVADCRGHLVWVGFSPDGSFLTS 900
 Db 841 STIQYCDFSPQNLAVVALSQYCVELWNTDSRSKVADCRGHLVWVGFSPDGSFLTS 900
 Qy 901 SDDQIRLWETKVKCKNSAVMLKQEVVDVVFQENWVWVLAVDHRRLOLINGRTGQIDYLT 960
 Db 901 SDDQIRLWETKVKCKNSAVMLKQEVVDVVFQENWVWVLAVDHRRLOLINGRTGQIDYLT 960
 Qy 961 EAQVSCCLSPHLOVIAFGDENGAIETLELVNRRIFQSRFOHKTVMHIOFTADEKTLIS 1020
 Db 961 EAQVSCCLSPHLOVIAFGDENGAIETLELVNRRIFQSRFOHKTVMHIOFTADEKTLIS 1020
 Qy 1021 SDDAEIOVWAKWOLDKCIPLRGHOETVKQFRLIKNSRLLSWSPGTQVKNVNIITGNKEK 1080
 Db 1021 SDDAEIOVWAKWOLDKCIPLRGHOETVKQFRLIKNSRLLSWSPGTQVKNVNIITGNKEK 1080
 Qy 1081 FVCHQGTVLSCDISHDATKFSSTSAKTAKIWSFDLLPLHELGHNGCVRCAPSVDS 1140
 Db 1081 FVCHQGTVLSCDISHDATKFSSTSAKTAKIWSFDLLPLHELGHNGCVRCAPSVDS 1140
 Qy 1141 LIATGDNGEIRIWNVSNGLLHLCAPLSEGAATHGGVNTDLCFSPDGKWLISAGGIYK 1200
 Db 1141 LIATGDNGEIRIWNVSNGLLHLCAPLSEGAATHGGVNTDLCFSPDGKWLISAGGIYK 1200
 Qy 1201 WNVVVTGESSQFTYNTGNLKKIHVSPDFKTYVVDNLGILYILOTLE 1248
 Db 1201 WNVVVTGESSQFTYNTGNLKKIHVSPDFKTYVVDNLGILYILOTLE 1248

AA97645
 ID AAY97645 standard; protein; 1248 AA.
 XX
 AC AAY97645;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Apaf-1XL-L83A protein sequence.
 XX
 KW Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 KW hereditary disease; Apaf-1XL-L83A.
 XX
 OS Homo sapiens.
 XX
 PN WO200100827-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 30-JUN-2000; 2000WO-US018039.
 XX
 PR 30-JUN-1999; 99US-0141718P.
 XX
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Nunez G, Hu Y;
 XX
 PP MPI; 2001-112454/12.
 XX
 DR N-PSDB; AAA91124.
 XX
 PT Compositions for screening apoptosis pathway agonists and antagonists,
 PT useful for the treatment and diagnosis of cancer and autoimmune diseases,
 PT comprises new splice variants of wild-type Apaf-1 gene.
 XX
 PS Disclosure; Fig 34; 101pp; English.
 XX
 CC This sequence represents an Apaf-1 variant of the invention. The variants
 CC of the invention are all splice variants of a wild-type Apaf-1 gene (a
 CC mammalian homolog of CED-4 that participates in cytochrome c-dependent
 CC activation of caspase-3). The DNA is useful in screening assays for
 CC identifying apoptosis signalling pathway (antagonists, which are in turn
 CC useful as potential therapeutics and diagnostic or prognostic tools for
 CC diverse types of cancers, autoimmune diseases and hereditary diseases and
 CC for screening compounds that modulates the interaction of Apaf-1 with
 CC other members of the signalling pathway, i.e. their substrates or
 CC ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 CC cell-free assay system and the antibody generated to the translation
 CC product are used in immunoprecipitation assays to isolate new Apaf-1
 CC pathway constituents or their natural mutants
 XX
 SQ Sequence 1248 AA;
 Query Match 99.9%; Score 6614; DB 4; Length 1248;
 Best Local Similarity 99.9%; Fred. No. 0;
 Matches 1247; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MDKARNCILLOHREALEKDIKTSYIMDMISDGLFTTSEEKURNEPTQOQRAAMLIYMI 60
 Db 1 MDKARNCILLOHREALEKDIKTSYIMDMISDGLFTTSEEKURNEPTQOQRAAMLIYMI 60
 Qy 61 LKKNDSYVSFYNNALHGEYKDLAALHDGIPVSSSSGKDSVSGITSYRTVLCGGVP 120
 Db 61 LKKNDSYVSFYNNALHGEYKDLAALHDGIPVSSSSGKDSVSGITSYRTVLCGGVP 120
 Qy 121 QRPVVFVTRKLVNAIQOKLSKQKCPGVWTHGMACGKSVLAABAVRDSHLLGCPFG 180
 Db 121 QRPVVFVTRKLVNAIQOKLSKQKCPGVWTHGMACGKSVLAABAVRDSHLLGCPFG 180
 Qy 181 GVHWVSQKQKSGGLMKLQNLCTRLDQDSFQRLPNTIEAKDRULMLRKHPSLL 240
 Db 181 GVHWVSQKQKSGGLMKLQNLCTRLDQDSFQRLPNTIEAKDRULMLRKHPSLL 240
 Qy 241 ILDDVWDSWVLKAFDQSQCIILLTRDKSVTDSVMGPKYVVPVSSLGKGLKLELSL 300

Db 241 ILDDVWDSWLVKAFDSQCIILLTRDKSVTVSVMGPKVVPVSSSLGKEGLEILSLFVN 300
 Qy 301 MKKADLPQAHSIIEKEGSPVWSLIGALLRDPFNWRYLLKQLNQKQKRIKSSSYD 360
 Db 301 MKKADLPQAHSIIEKEGSPVWSLIGALLRDPFNWRYLLKQLNQKQKRIKSSSYD 360
 Qy 361 YEALDEAMSIIVEMLRDIDKYITDLSILQKDVKPTKVLCLMDMETERVEDILOEFVN 420
 Db 361 YEALDEAMSIIVEMLRDIDKYITDLSILQKDVKPTKVLCLMDMETERVEDILOEFVN 420
 Qy 421 KSLIFCDNRGKSFRIYLDLQVDFTEKNCQQLDHLKKIITQFORHOPHTLSPDQDC 480
 Db 421 KSLIFCDNRGKSFRIYLDLQVDFTEKNCQQLDHLKKIITQFORHOPHTLSPDQDC 480
 Qy 481 MYWYNFLAYHMASAKMHELCALMFLDWIKAKTELVGPAHLIHEFVEYRHLDEKCAV 540
 Db 481 MYWYNFLAYHMASAKMHELCALMFLDWIKAKTELVGPAHLIHEFVEYRHLDEKCAV 540
 Qy 541 SENTQEFSLNGHLLGROPPFNIIVOLGLCEBETSEVYQOAKLOAKQEVNDGMLYLEWINK 600
 Db 541 SENTQEFSLNGHLLGROPPFNIIVOLGLCEBETSEVYQOAKLOAKQEVNDGMLYLEWINK 600
 Qy 601 KNIITNLSRLVVRPHDTAVYHACFSDGRIASCGADKTLQVKAETGEKLEIKAHEDV 660
 Db 601 KNIITNLSRLVVRPHDTAVYHACFSDGRIASCGADKTLQVKAETGEKLEIKAHEDV 660
 Qy 661 LCCAFSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
 Db 661 LCCAFSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
 Qy 721 SDCLFLKLDLWLNQKCRNTMFGHTSNVHCRSPDDKLLASCSADGTLKLWDATSANERKS 780
 Db 721 SDCLFLKLDLWLNQKCRNTMFGHTSNVHCRSPDDKLLASCSADGTLKLWDATSANERKS 780
 Qy 781 INVQKFFLNLDPOEDMEVIVKCCSWASADGARIMVAANKIIFLDIHTSGLLGIHTGHH 840
 Db 781 INVQKFFLNLDPOEDMEVIVKCCSWASADGARIMVAANKIIFLDIHTSGLLGIHTGHH 840
 Qy 841 STIQYCDSPQNHVALVVALSQCVELWNTDSRSKVDRCRHLVHGVWVSPDGSSFLTS 900
 Db 841 STIQYCDSPQNHVALVVALSQCVELWNTDSRSKVDRCRHLVHGVWVSPDGSSFLTS 900
 Qy 901 SDDQIRLWETKVKCNKSNVLMKQEVVVFQENEVMLAVDHIIRLQLINGRTQIDYLT 960
 Db 901 SDDQIRLWETKVKCNKSNVLMKQEVVVFQENEVMLAVDHIIRLQLINGRTQIDYLT 960
 Qy 961 EAOVSCCLSPHLOVIAFGDENGAIETILELVNNRIFQSRFOHKKTVWHIQTADKTLIS 1020
 Db 961 EAOVSCCLSPHLOVIAFGDENGAIETILELVNNRIFQSRFOHKKTVWHIQTADKTLIS 1020
 Qy 1021 SSDDAEIQVNNWQDKCIFLFGHQBTVKDFRLLKNSRLLSWDFGTVKVWNIITGNKEKD 1080
 Db 1021 SSDDAEIQVNNWQDKCIFLFGHQBTVKDFRLLKNSRLLSWDFGTVKVWNIITGNKEKD 1080
 Qy 1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLPLHELGHNGCVRCSPAFSDST 1140
 Db 1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLPLHELGHNGCVRCSPAFSDST 1140
 Qy 1141 LLATGDDNGEIRIWNVSGELLHLICAPLSEGAATHGCGWTVDLFCSPDGKMLISAGGYIK 1200
 Db 1141 LLATGDDNGEIRIWNVSGELLHLICAPLSEGAATHGCGWTVDLFCSPDGKMLISAGGYIK 1200
 Qy 1201 WNNVVTGESSQTFYTNGLNKKIHVSDFKTYVTVDNLGILYILOTLE 1248
 Db 1201 WNNVVTGESSQTFYTNGLNKKIHVSDFKTYVTVDNLGILYILOTLE 1248

RESULT 9
 ID AA97641
 standard; protein; 1248 AA.
 AC AA97641;

XX 20-APR-2001 (first entry)
 XX Apaf-1XL-L10A protein sequence.
 DE DE Apaf1: CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 KW hereditary disease; Apaf-1XL-L10A.
 XX Homo sapiens.
 XX WO200100827-A1.
 XX 04-JAN-2001.
 XX 30-JUN-2000; 2000WO-US018039.
 XX 30-JUN-1999; 99US-0141718P.
 XX (UNMI) UNIV MICHIGAN.
 PA Nunez G, Hu Y;
 XX MPI; 2001-112454/12.
 XX N-PSDB; AAA91120.
 XX Compositions for screening apoptosis pathway agonists and antagonists,
 XX useful for the treatment and diagnosis of cancer and autoimmune diseases,
 XX comprises new splice variants of wild-type Apaf-1 gene.
 XX Disclosure; Fig 30; 101pp; English.
 XX This sequence represents an Apaf-1 variant of the invention. The variants
 CC of the invention are all splice variants of a wild-type Apaf-1 gene (a
 CC mammalian homolog of CED-4 that participates in cytochrome c-dependent
 CC activation of caspase-3). The DNA is useful in screening assays for
 CC identifying apoptosis signalling pathway (antagonists, which are in turn
 CC useful as potential therapeutics and diagnostic or prognostic tools for
 CC diverse types of cancers, autoimmune diseases and hereditary diseases and
 CC for screening compounds that modulates the interaction of Apaf-1 with
 CC other members of the signalling pathway, i.e. their substrates or
 CC ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 CC cell-free assay system and the antibody generated to the translation
 CC product are used in immunoprecipitation assays to isolate new Apaf-1
 CC pathway constituents or their natural mutants

Sequence 1248 AA;

Query Match 99.9%; Score 6614; DB 4; Length 1248;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1247; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MDKARVCLLQNRALKEDIKTSYIMDHMSDGLTISEEEKVNEPTQOQRAAMLIKMI 60
 Db 1 MDKARNVCLQNRALKEDIKTSYIMDHMSDGLTISEEEKVNEPTQOQRAAMLIKMI 60
 Qy 61 LKQNDSDVSVFYNNALLHEGYKDLAALLHDGIPVVSSSSGKDSVSGITSVYRTVLCGGVP 120
 Db 61 LKQNDSDVSVFYNNALLHEGYKDLAALLHDGIPVVSSSSGKDSVSGITSVYRTVLCGGVP 120
 Qy 121 QRPVVFTRKKLVNAIQKLSKLGEPGWVTHGMAGCGKSVLAABAVRDSHSLLEGCPGP 180
 Db 121 QRPVVFTRKKLVNAIQKLSKLGEPGWVTHGMAGCGKSVLAABAVRDSHSLLEGCPGP 180
 Qy 181 GVHWVSVGKODKSGLLMKLQNLCTRLDQDSFSORLPNTEEAKDRILMLRKHPSLL 240
 Db 181 GVHWVSVGKODKSGLLMKLQNLCTRLDQDSFSORLPNTEEAKDRILMLRKHPSLL 240
 Qy 241 ILDDVWDSWLVKAFDSQCIILLTRDKSVTVSVMGPKVVPVSSSLGKEGLEILSLFVN 300
 Db 241 ILDDVWDSWLVKAFDSQCIILLTRDKSVTVSVMGPKVVPVSSSLGKEGLEILSLFVN 300
 Qy 301 MKKADLPQAHSIIEKEGSPVWSLIGALLRDPFNWRYLLKQLNQKQKRIKSSSYD 360

Db 301 MKKADLPEQAHSIIKECKGSLVLSLIGALLRDFPNRWEYLLQKQKFRIRKSSSYD 360
 Qy 361 YEALDEAMSISVEMLRDEIKDYITDLSILOQKQKVPKVLCTILMDMETEEVEDILOEFVN 420
 Db 361 YEALDEAMSISVEMLRDEIKDYITDLSILOQKQKVPKVLCTILMDMETEEVEDILOEFVN 420
 Qy 421 KSLIFCDNRNGKFRYYLHDLQVDFLTCKNGSQDLHKKIITQFORVHQPHTLSPDOEDC 480
 Db 421 KSLIFCDNRNGKFRYYLHDLQVDFLTCKNGSQDLHKKIITQFORVHQPHTLSPDOEDC 480
 Qy 481 MYWYNFLAYHMASAKWHELCALMFSLDWKATKELVGPALHIEFEYRHIIDKDCAV 540
 Db 481 MYWYNFLAYHMASAKWHELCALMFSLDWKATKELVGPALHIEFEYRHIIDKDCAV 540
 Qy 541 SENFQEFSLNGHLGROPPFNIVQLGCEPTESEVYQAKLOKQEVNDGMLYLEWINK 600
 Db 541 SENFQEFSLNGHLGROPPFNIVQLGCEPTESEVYQAKLOKQEVNDGMLYLEWINK 600
 Qy 601 KNITNLSRLVVRPHTDAVYHACFSEDOQRIASCGADKTLQVFKATGEKLEIKAHEDV 660
 Db 601 KNITNLSRLVVRPHTDAVYHACFSEDOQRIASCGADKTLQVFKATGEKLEIKAHEDV 660
 Qy 661 LCCAFSTDDRFATCSVDKKVIMNSMTGELVHTYDESEQVNCCHFTNSHLLATGS 720
 Db 661 LCCAFSTDDRFATCSVDKKVIMNSMTGELVHTYDESEQVNCCHFTNSHLLATGS 720
 Qy 721 SDCFLKLDLNOKECRNFMGHTNSVNHCRFSPDDKLLASCSADGTLKLDATSANERKS 780
 Db 721 SDCFLKLDLNOKECRNFMGHTNSVNHCRFSPDDKLLASCSADGTLKLDATSANERKS 780
 Qy 781 INVQPFLEDPQDEMEVIVKCCMSADGARIMAAKNKIPFLDIHTSGLLGEIHTGHH 840
 Db 781 INVQPFLEDPQDEMEVIVKCCMSADGARIMAAKNKIPFLDIHTSGLLGEIHTGHH 840
 Qy 841 STIQYCDSPQNHLLAVALSQYVELWNTDSRSKVADCRGHLVSWHGVMSFGSSSFLTS 900
 Db 841 STIQYCDSPQNHLLAVALSQYVELWNTDSRSKVADCRGHLVSWHGVMSFGSSSFLTS 900
 Qy 901 SDDQITRLWETKVKCKNSAVMLKQEVNDVVFQENEVMVLAVDHIRRLQIINGRTGQIDYLT 960
 Db 901 SDDQITRLWETKVKCKNSAVMLKQEVNDVVFQENEVMVLAVDHIRRLQIINGRTGQIDYLT 960
 Qy 961 EAQVSCCLSLPHLOIYAFGDENGAIETLELVNRRIFQSRFOHKKTVWHIOFTADEKTLIS 1020
 Db 961 EAQVSCCLSLPHLOIYAFGDENGAIETLELVNRRIFQSRFOHKKTVWHIOFTADEKTLIS 1020
 Qy 1021 SSDDAEIQVNNWQDKCIFLRGHOETVKOPRLLKNSRLLSWSFDGTWKVWNIIITGNKXD 1080
 Db 1021 SSDDAEIQVNNWQDKCIFLRGHOETVKOPRLLKNSRLLSWSFDGTWKVWNIIITGNKXD 1080
 Qy 1081 FVCHQGTVLSCDISHDATKFSSTADKTIKINSFDLILPLHELGHNGCVRCASFVDS 1140
 Db 1081 FVCHQGTVLSCDISHDATKFSSTADKTIKINSFDLILPLHELGHNGCVRCASFVDS 1140
 Qy 1141 LLATGDDNGEIRINWVNSGELLHLHCAPLSEGAATHGQVTDLCFSPDGKMLISAGGYIK 1200
 Db 1141 LLATGDDNGEIRINWVNSGELLHLHCAPLSEGAATHGQVTDLCFSPDGKMLISAGGYIK 1200
 Qy 1201 WNNVVTGESSQFTYNGTNLKKIHVSPDKTYVTVDNLGILYILOTLE 1248
 Db 1201 WNNVVTGESSQFTYNGTNLKKIHVSPDKTYVTVDNLGILYILOTLE 1248

RESULT 10
 ID AAY97647
 AC AAY97647;
 XX AAY97647;
 XX 20-APR-2001 (first entry)
 DT
 XX
 DE Apaf-1XL-M368L/K160R protein sequence.

XX Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 KW hereditary disease; Apaf-1XL-M368L/K160R.
 XX Homo sapiens.
 XX WC200100827-A1.
 XX 04-JAN-2001.
 XX 30-JUN-2000; 2000WO-US018039.
 XX 30-JUN-1999; 99US-0141718P.
 XX (UNMI) UNIV MICHIGAN.
 XX Nunez G, Hu Y;
 XX WPI: 2001-112454/12.
 XX N-PSDB; AAN91126.
 XX Compositions for screening apoptosis pathway agonists and antagonists,
 useful for the treatment and diagnosis of cancer and autoimmune diseases,
 comprises new splice variants of wild-type Apaf-1 Gene.
 XX Disclosure; Fig 36; 101pp; English.
 XX This sequence represents an Apaf-1 variant of the invention. The variants
 of the invention are all splice variants of a wild-type Apaf-1 gene (a
 mammalian homolog of CED-4 that participates in cytochrome c-dependent
 activation of caspase-3). The DNA is useful in screening assays for
 identifying apoptosis signalling pathway (antagonists, which are in turn
 useful as potential therapeutic and diagnostic or prognostic tools for
 diverse types of cancers, autoimmune diseases and hereditary diseases and
 for screening compounds that modulates the interaction of Apaf-1 with
 other members of the signalling pathway, i.e. their substrates or
 ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 cell-free assay system and the antibody generated to the translation
 product are used in immunoprecipitation assays to isolate new Apaf-1
 pathway constituents or their natural mutants

Query Match 99.9%; Score 6613; DB 4; Length 1248;
 Best Local Similarity 99.8%; Pred No. 0;
 Matches 1246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDKARNCLLOHREALEKDIKTSYIMDHWSIDGFLTISEEKVNEPTQOQRAAMLIKMI 60
 Db 1 MDKARNCLLOHREALEKDIKTSYIMDHWSIDGFLTISEEKVNEPTQOQRAAMLIKMI 60
 Qy 61 LKKNDSYVSYFYNALLHREGYKDLAALLHDGIPVYSSSGKDSVSGITSYVTVLCEGVP 120
 Db 61 LKKNDSYVSYFYNALLHREGYKDLAALLHDGIPVYSSSGKDSVSGITSYVTVLCEGVP 120
 Qy 121 QRPVVFTRKKLVNAIQOKLSKLGEPGWVTHGMAGGKSVLAABAVRDHSLLEGCPFG 180
 Db 121 QRPVVFTRKKLVNAIQOKLSKLGEPGWVTHGMAGGKSVLAABAVRDHSLLEGCPFG 180
 Qy 181 GVHWVSGKQDKSGLLMKLNQCTLDQDESFSORPLNTEEAKDLRLILMLRKHPRSL 240
 Db 181 GVHWVSGKQDKSGLLMKLNQCTLDQDESFSORPLNTEEAKDLRLILMLRKHPRSL 240
 Qy 241 ILDDVDSWLVKAFDSQCCQILITTRDKSVTDSVNGPKYVYVVPVSSSLGKGLIILSFVN 300
 Db 241 ILDDVDSWLVKAFDSQCCQILITTRDKSVTDSVNGPKYVYVVPVSSSLGKGLIILSFVN 300
 Qy 301 MKKADLPEQAHSIIKECKGSLVLSLIGALLRDFPNRWEYLLQKQKFRIRKSSSYD 360
 Db 301 MKKADLPEQAHSIIKECKGSLVLSLIGALLRDFPNRWEYLLQKQKFRIRKSSSYD 360
 Qy 361 YEALDEAMSISVEMLRDEIKDYITDLSILOQKQKVPKVLCTILMDMETEEVEDILOEFVN 420

Db 470 MYTNFLAYHNASAKQHKELCALMESLDWIIKAKTELVGPAHLIHEFVYRHLDEKCAV 529
Qy 541 SENFQEFSLNGLHGRFPFNIVOLGCEPETSEVYQAKLQAKQEVNDGMLYLEWINK 600
Db 530 SENFQEFSLNGLHGRFPFNIVOLGCEPETSEVYQAKLQAKQEVNDGMLYLEWINK 589
Qy 601 KNIITNLSRLVVRPHTDAVYHACFSEDDGQRIASCAGADKTLQVFKAPTGBKLLEIKAHEDV 660
Db 590 KNIITNLSRLVVRPHTDAVYHACFSEDDGQRIASCAGADKTLQVFKAPTGBKLLEIKAHEDV 649
Qy 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEOVNCCHTNSHHLLLATGS 720
Db 650 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEOVNCCHTNSHHLLLATGS 709
Qy 721 SDCFLKLDLWLNQKCEKRNMTFHTSNVHCRFSPDDKLLASCADGTLKLDATSANERKS 780
Db 710 SDCFLKLDLWLNQKCEKRNMTFHTSNVHCRFSPDDKLLASCADGTLKLDATSANERKS 769
Qy 781 INVQOFFLNLDEPQEDMEVIVKCCSWASADGARIMVAANKIFLFDIHTSGLLGEIHTGHH 840
Db 770 INVQOFFLNLDEPQEDMEVIVKCCSWASADGARIMVAANKIFLFDIHTSGLLGEIHTGHH 829
Qy 841 STIOYCFSPONHLAVVALSOYCVELWNTDSRKVADCRGHLVWGVMPSPDGSSFLTS 900
Db 830 STIOYCFSPONHLAVVALSOYCVELWNTDSRKVADCRGHLVWGVMPSPDGSSFLTS 889
Qy 901 SDDQIRLWETKVKCKSNVLMKQEVVVOFENEVWVLAVDHIRRLQNLINGRTGQIDYLT 960
Db 890 SDDQIRLWETKVKCKSNVLMKQEVVVOFENEVWVLAVDHIRRLQNLINGRTGQIDYLT 949
Qy 961 EAQVSCCCLSPHLOVAFDENGAEIILELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 1020
Db 950 EAQVSCCCLSPHLOVAFDENGAEIILELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 1009
Qy 1021 SSDAEIQVWQWLDKCIPLRGHQTVDKDFLLKXNRLLSWSPDGTVKVWNIITGNKEKD 1080
Db 1010 SSDAEIQVWQWLDKCIPLRGHQTVDKDFLLKXNRLLSWSPDGTVKVWNIITGNKEKD 1069
Qy 1081 FVCHQGTVLSCDISHDAPKFSSTADKTAKTWSFDLLPLHLRGHNGCVCSAFSDVST 1140
Db 1070 FVCHQGTVLSCDISHDAPKFSSTADKTAKTWSFDLLPLHLRGHNGCVCSAFSDVST 1129
Qy 1141 LLATGDDNGEIRIWNVNSGELLHLICAPLSEGAATHGQWVTDLCFSPDGKMLISAGGYIK 1200
Db 1130 LLATGDDNGEIRIWNVNSGELLHLICAPLSEGAATHGQWVTDLCFSPDGKMLISAGGYIK 1189
Qy 1201 WNVVVTGESSQTFYTNCTNLKIHVSPDKTYVTVDNLGILYILQTL 1248
Db 1190 WNVVVTGESSQTFYTNCTNLKIHVSPDKTYVTVDNLGILYILQTL 1237
RESULT 13
ADP65308 standard; protein; 1237 AA.
XX AC ADP65308;
XX DT 12-AUG-2004 (first entry)
XX DE Human apoptotic protease activating factor isoform A, apoptotic protease.
XX KW autoimmune disease; arthritis; gene expression analysis;
KW KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;
KW KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
KW KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
XX KW immune; human.
OS Homo sapiens.
XX XX
FN W02003072827-A1.
XX XX

PD 04-SEP-2003.
XX 31-OCT-2002; 2002WO-US035433.
PF 31-OCT-2001; 2001US-0336220P.
PR (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
PA Hirsch R, Thorton SL;
XX WPI; 2003-712740/67.
XX GENBANK; NP_037361.
PT Diagnosing and analyzing autoimmune disease using gene expression
PT profiles and microarray technology, useful for diagnosing and treating
PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
PT gout.
XX Disclosure; Page; 56pp; English.
XX The invention relates to a novel method for diagnosing and analysing
CC autoimmune disease or arthritides. The method comprises obtaining a
CC patient sample containing mRNA, analysing gene expression using the mRNA
CC that results in a gene expression signature of the mRNA, and using that
CC gene expression signature to diagnose or analyse the autoimmune disease
CC or arthritides in the patient, where gene expression of at least 60% of
CC the genes correlates with that of the gene signature. The invention
CC further comprises: a treatment of rheumatoid arthritis; identification of
CC genes for targeting in the treatment of rheumatoid arthritis in a mammal
CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
CC array or gene chip, specific for rheumatoid arthritis; diagnosis or
CC analyses of autoimmune disease or rheumatoid arthritis; screening the
CC efficacy of a candidate drug in vitro for the treatment of collagen-
CC induced arthritis; and reducing the symptoms associated with collagen-
CC induced arthritis. The compositions of the invention have the following
CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
CC antigout, antiinflammatory, dermatological, and immunomodulatory. The
CC methods and compositions of the present invention are useful for
CC diagnosing and treating autoimmune disease or arthritides, such as
CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis, and
CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
CC immune disease caused by an infectious agent. This sequence represents a
CC protein sequence relating to the genes used in the analysis and treatment
CC of autoimmune diseases or arthritides. Note: This sequence is not shown
CC in the specification. It has been supplied in an electronic format from
CC WIPO.
XX SQ Sequence 1237 AA;

Query Match 98.9%; Score 6548.5; DB 7; Length 1237;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1237; Conservative 0; Mismatches 0; Indels 11; Gaps 1;
Qy 1 MDAKARNCLLOHREALEKDIKTSYIMDMHISDGLFTISEEEKVNEPTQOQRAAMLIKMI 60
Db 1 MDAKARNCLLOHREALEKDIKTSYIMDMHISDGLFTISEEEKVNEPTQOQRAAMLIKMI 60
Qy 61 LKDNDSYVSFYNNALLHEGYKDLAALLHDGTPVYSSSSGKDSVSGITSYVTVLCEGGVP 120
Db 61 LKDNDSYVSFYNNALLHEGYKDLAALLHDGTPVYSSSSGKDSVSGITSYVTVLCEGGVP 109
Qy 121 QRPVVFVTRKLVNVAIQKLSKLBEGFGWVTHGMAGCKSVLAEEAVRDHSLLEGCPFG 180
Db 110 QRPVVFVTRKLVNVAIQKLSKLBEGFGWVTHGMAGCKSVLAEEAVRDHSLLEGCPFG 169
Qy 181 GVHWVSVGKDQKSGLLMKLQNLCTRLQDQDSFSORLPINTEAKDRILMLRKHPRSL 240
Db 170 GVHWVSVGKDQKSGLLMKLQNLCTRLQDQDSFSORLPINTEAKDRILMLRKHPRSL 229
Qy 241 ILDDVWDSWVLKAFDSCQILLTTRDKSVTDSVNGPKYVVPVSSLGKGLTILSLFVN 300
Db 230 ILDDVWDSWVLKAFDSCQILLTTRDKSVTDSVNGPKYVVPVSSLGKGLTILSLFVN 289

QY 301 MKKADLPQAHSHKECKGSPVLSLIGALLDFPNRWEYVLKOLQKQPKRIRKSSSYD 360
 Db 290 MKKADLPQAHSHKECKGSPVLSLIGALLDFPNRWEYVLKOLQKQPKRIRKSSSYD 349
 QY 361 YEALDEAMSISVEMLRDIDKYDTLSDLSILQKQKVPKVLKILWDMETEVEDILQEFVN 420
 Db 350 YEALDEAMSISVEMLRDIDKYDTLSDLSILQKQKVPKVLKILWDMETEVEDILQEFVN 409
 QY 421 KSLAFCDRNGKSPRYLLHDLQVDFLTKNCSQLODLHKHIIQFORVHQPHTLSPDOEDC 480
 Db 410 KSLAFCDRNGKSPRYLLHDLQVDFLTKNCSQLODLHKHIIQFORVHQPHTLSPDOEDC 469
 QY 481 MYWYNFLAYHWSAKMKEKCALMFSLDKAKTELUGPAHLIHEFVEVHILDEKCAV 540
 Db 470 MYWYNFLAYHWSAKMKEKCALMFSLDKAKTELUGPAHLIHEFVEVHILDEKCAV 529
 QY 541 SENFOEFLSLNGHLLGPPNNIVQLGCEPETSVEYQAKLOAKQVDMGMLYLEWINK 600
 Db 530 SENFOEFLSLNGHLLGPPNNIVQLGCEPETSVEYQAKLOAKQVDMGMLYLEWINK 589
 QY 601 KNITNLSRLVVRPHTDAVYHACFSESDGQRIASCGADKTLQVKAETGEKLEIKAHEDV 660
 Db 590 KNITNLSRLVVRPHTDAVYHACFSESDGQRIASCGADKTLQVKAETGEKLEIKAHEDV 649
 QY 661 LCCAFSTDDREIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLLATGS 720
 Db 650 LCCAFSTDDREIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLLATGS 709
 QY 721 SDCFLKLDLNOKECRNMTGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
 Db 710 SDCFLKLDLNOKECRNMTGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 769
 QY 781 INVKOFFNLNEDPOEDMEVIVKCCSADGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 840
 Db 770 INVKOFFNLNEDPOEDMEVIVKCCSADGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 829
 QY 841 STIQVCDSPQNHLLAVLALSOYCVELWNTDSKVDACRGLHSMVHGVMSPDGSSFLTS 900
 Db 830 STIQVCDSPQNHLLAVLALSOYCVELWNTDSKVDACRGLHSMVHGVMSPDGSSFLTS 889
 QY 901 SDOQIRLWETKVKCKNSAVMLKQEDVVVFOENWVLAVDHIRELOLINGRTGQIDYLT 960
 Db 890 SDOQIRLWETKVKCKNSAVMLKQEDVVVFOENWVLAVDHIRELOLINGRTGQIDYLT 949
 QY 961 EAQVSCCLSPHLQVIAFGDENGAIIELELVNRRIFQSRFOHKTVMWHIQTADKTLIS 1020
 Db 950 EAQVSCCLSPHLQVIAFGDENGAIIELELVNRRIFQSRFOHKTVMWHIQTADKTLIS 1009
 QY 1021 SSDDAEIQVNWQDLKCIFLRGHQTQKDFRLIKNSRLLSWSFDGTVKVWNIIITGNKEKD 1080
 Db 1010 SSDDAEIQVNWQDLKCIFLRGHQTQKDFRLIKNSRLLSWSFDGTVKVWNIIITGNKEKD 1069
 QY 1081 FVCHQGTVLSCDISHDATKFSSTADKTKIWSFDLLPLHELGRHNGCVRCSAFVSVDST 1140
 Db 1070 FVCHQGTVLSCDISHDATKFSSTADKTKIWSFDLLPLHELGRHNGCVRCSAFVSVDST 1129
 QY 1141 LLATGDDNGEIRIWNVNGELHLHLCAPLSEGAATHGWWTDLCFSPDGKMLISAGGIYK 1200
 Db 1130 LLATGDDNGEIRIWNVNGELHLHLCAPLSEGAATHGWWTDLCFSPDGKMLISAGGIYK 1189
 QY 1201 WNVVGTSSQTFVNGTNLKKHVSFDPKTVYVNDNLGILYILOTLE 1248
 Db 1190 WNVVGTSSQTFVNGTNLKKHVSFDPKTVYVNDNLGILYILOTLE 1237

RESULT 14
 AAW91072
 ID AAW91072 standard; protein; 1205 AA.
 XX
 AC AAW91072;
 XX
 DT 07-JUN-1999 (first entry)
 XX

DE Apoptosis inducer splice variant Apaf-1L.
 XX Apaf-1L; Apaf-1; splice variant; human; caspase-3; activator; Ced-4;
 KW human; apoptosis; programmed cell death; cancer; AIDS;
 KW multiple sclerosis; inflammation; therapy.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 99..109
 FT Region /note= "additional 11 amino acids compared with Apaf-1"
 FT 613..643
 FT /note= "putative WD repeat"
 FT 655..685
 FT /note= "putative WD repeat"
 FT 692..729
 FT /note= "putative WD repeat"
 FT 741..771
 FT /note= "putative WD repeat"
 FT 788..825
 FT /note= "putative WD repeat"
 FT 928..946
 FT /note= "putative WD repeat"
 FT 958..988
 FT /note= "putative WD repeat"
 FT 999..1028
 FT /note= "putative WD repeat"
 FT 1040..1070
 FT /note= "putative WD repeat"
 FT 1082..1112
 FT /note= "putative WD repeat"
 FT 1119..1160
 FT /note= "putative WD repeat"
 XX
 XX WO9855615-A1.
 XX 10-DEC-1998.
 XX 05-JUN-1998; 98WO-US011773.
 XX 05-JUN-1997; 97US-0048807P.
 XX 07-AUG-1997; 97US-0055258P.
 XX (GETH) GENENTECH INC.
 XX (TEXA) UNIV TEXAS.
 XX Zou H, Henzel WJ, Wang X;
 XX WPI; 1999-080828/07.
 XX N-PSDB; AAV84798.
 XX
 XX New Apaf-1 protein for regulating apoptosis - is a human homologue of Ced
 XX -4 and an activator of caspase-3; for screening agents that regulate
 XX apoptosis and for treating cancer, AIDS.
 XX Claim 4; Fig 17; 134pp; English.
 XX
 XX This polypeptide comprises Apaf-1L, a splice variant of Apaf-1 (see
 XX AAW91071), a novel human protein that is believed to be a mammalian
 XX homologue of the Caenorhabditis elegans Ced-4 apoptosis protein. Apaf-1L
 XX cDNA (see AAV84798) was produced from HeLa cell mRNA. The encoded
 XX polypeptide contains 11 additional amino acids compared to Apaf-1, but
 XX retains the structural and functional features of Apaf-1. It is a 130 kDa
 XX polypeptide containing a Ced-3 homologous N-terminal domain, followed by
 XX a Ced-4 homologous domain and multiple C-terminal WD-40 repeats. It forms
 XX a complex with cytochrome-3 and activates caspase-3 in the apoptotic
 XX pathway. Apaf-1L can be produced by culturing cells transformed or
 XX transfected with a vector containing Apaf-1L nucleic acid. Apaf-1
 XX polypeptides, including Apaf-1L, can be used to induce apoptosis, and the
 XX polypeptides, genes and antibodies are especially useful for screening
 XX and identifying cells signalled for apoptosis. Agonistic antibodies can
 XX be used to stimulate apoptosis in cancer cells, and antagonistic
 XX antibodies can be used to block excessive apoptosis or to block the

CC autoimmune and antiinflammatory activities of Apaf-1, in conditions such
 CC as AIDS and multiple sclerosis. Apaf-1 polypeptides and nucleic acids can
 CC be used to screen for agents which modulate apoptosis. Knockout animals
 CC are also provided
 XX
 SQ Sequence 1205 AA;

Query Match 96.0%; Score 6352.5; DB 2; Length 1205;
 Best Local Similarity 96.6%; Pred. No. 0;
 Matches 1205; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

Qy	1	MDAKARNCLLOHREALEXDKITSYIMDHMSDGLTISEEEKVRNEPTQOQRAAMLIKMI	60
Db	1	MDAKARNCLLOHREALEXDKITSYIMDHMSDGLTISEEEKVRNEPTQOQRAAMLIKMI	60
Qy	61	LKKONDSYVSFYNNALHGGYKDLAALLHDGIPVSSSGKDSVSGITSYVTVLCEGGVP	120
Db	61	LKKONDSYVSFYNNALHGGYKDLAALLHDGIPVSSSGKDSVSGITSYVTVLCEGGVP	120
Qy	121	QRPVVVTRKKLVNAIQKLSKLKGEPEGWTHHGAGCKSVLAABAVRDHSLLEGCPFG	180
Db	121	QRPVVVTRKKLVNAIQKLSKLKGEPEGWTHHGAGCKSVLAABAVRDHSLLEGCPFG	180
Qy	181	GVHWVSVCKQDKSGLLMLKQLNCTELDODESQRPLNIEBAKDLRLMLRKHPRSLL	240
Db	181	GVHWVSVCKQDKSGLLMLKQLNCTELDODESQRPLNIEBAKDLRLMLRKHPRSLL	240
Qy	241	ILDDVWDGWLKAFDSQOQILTTTRDKSVTSVMGPKVVPVSSSLGKEGLEILSLFN	300
Db	241	ILDDVWDGWLKAFDSQOQILTTTRDKSVTSVMGPKVVPVSSSLGKEGLEILSLFN	300
Qy	301	MKADLPQOAHSIIEKEGSPPLVSLIGALLRDPFNWYVYLKQLNQKPKRIRKSSSYD	360
Db	301	MKADLPQOAHSIIEKEGSPPLVSLIGALLRDPFNWYVYLKQLNQKPKRIRKSSSYD	360
Qy	361	YEALDEAMSIIVEMUREDIKDYITDLSILOKDKVPTKVLCLMDMETEEVEDILOQFVN	420
Db	361	YEALDEAMSIIVEMUREDIKDYITDLSILOKDKVPTKVLCLMDMETEEVEDILOQFVN	420
Qy	421	KSLFPCDRNGKFRYYLHDQVDFTEKNCQQLDHLKKITQFORYHQPHTLSPDQDC	480
Db	421	KSLFPCDRNGKFRYYLHDQVDFTEKNCQQLDHLKKITQFORYHQPHTLSPDQDC	480
Qy	481	MYWNFLAYHMASAKWHKELCALMFSLDWIKAKTELVGPAHLIHEFVEYRHLDEKCAV	540
Db	481	MYWNFLAYHMASAKWHKELCALMFSLDWIKAKTELVGPAHLIHEFVEYRHLDEKCAV	540
Qy	541	SENFOEFLSLNGHLLGROPFPFNIVOLGLCEPETSVEVYQOAKLQAKQEVNDGMLYLEWINK	600
Db	541	SENFOEFLSLNGHLLGROPFPFNIVOLGLCEPETSVEVYQOAKLQAKQEVNDGMLYLEWINK	600
Qy	601	KNITNLSRLVVRPHPTDVAHYHACFSDGQRIASCADKTLQVFKAETGKLEIKAHEDEV	660
Db	601	KNITNLSRLVVRPHPTDVAHYHACFSDGQRIASCADKTLQVFKAETGKLEIKAHEDEV	660
Qy	661	LCCAFSTDDRFRATCSVDKVKIWNMTGELVHTYDEHSEVNCCHFTNSHHLLATGS	720
Db	661	LCCAFSTDDRFRATCSVDKVKIWNMTGELVHTYDEHSEVNCCHFTNSHHLLATGS	720
Qy	721	SDCFLLKLDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS	780
Db	721	SDCFLLKLDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS	780
Qy	781	INVKOFFLNLNDEPOEDMEVIVKCCSWSDAGRIWAAKNKIFLFDIHTSGLLGEIHTGHH	840
Db	781	INVKOFFLNLNDEPOEDMEVIVKCCSWSDAGRIWAAKNKIFLFDIHTSGLLGEIHTGHH	840
Qy	841	STIQYCDSPQNHVALVVALSQVCELWNTDGRSKVADCRGHLWSVHWGFMFPGDSSFLTS	900
Db	841	STIQYCDSPQNHVALVVALSQVCELWNTDGRSKVADCRGHLWSVHWGFMFPGDSSFLTS	900
Qy	901	SDDOTIRLWETKVKCKNSAVMLKQEVVDVVFQENVMVLAVDHRLQLINGRTGQIDYLT	960
Db	901	SDDOTIRLWETKVKCKNSAVMLKQEVVDVVFQENVMVLAVDHRLQLINGRTGQIDYLT	960

Db	858	SDDOTIRLWETKVKCKNSAVMLKQEVVDVVFQENVMVLAVDHRLQLINGRTGQIDYLT	917
Qy	961	EAQVSCCLSPHLQYIAFGDENGAIIEILELVNNRIFQSRFQHKKTVMHIQFTADEKTLIS	1020
Db	918	EAQVSCCLSPHLQYIAFGDENGAIIEILELVNNRIFQSRFQHKKTVMHIQFTADEKTLIS	977
Qy	1021	SSDDAEIOVWVWQDCKCIFLRGHQETVKDFRLLKNSRLLSWSFDGTVKVMNIITGNKEKD	1080
Db	978	SSDDAEIOVWVWQDCKCIFLRGHQETVKDFRLLKNSRLLSWSFDGTVKVMNIITGNKEKD	1037
Qy	1081	FVCHQGTVLSGDISHDATKFSSTADKTAKTWSFDLPLPLHLRGHNGCVRCSAFSVOST	1140
Db	1038	FVCHQGTVLSGDISHDATKFSSTADKTAKTWSFDLPLPLHLRGHNGCVRCSAFSVOST	1097
Qy	1141	LLATGDDNGEIRIWNVNGELHLHLCAPLSEGAATHGCGWVTDLCFSPDGKMLISAGGIK	1200
Db	1098	LLATGDDNGEIRIWNVNGELHLHLCAPLSEGAATHGCGWVTDLCFSPDGKMLISAGGIK	1157
Qy	1201	WNVVVTGESSQTFYNTGNTLNKKIHSVDPFKTYVTVVDNLGILYLQTL	1248
Db	1158	WNVVVTGESSQTFYNTGNTLNKKIHSVDPFKTYVTVVDNLGILYLQTL	1205

RESULT 15
 AAY97637
 ID AAY97637 standard; protein; 1205 AA.
 XX
 AC AAY97637;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Apaf-1L protein sequence.
 XX
 KW Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 KW hereditary disease; Apaf-1L.
 XX
 OS Homo sapiens.
 XX
 PN WO200100827-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 30-JUN-2000; 2000WO-US018039.
 XX
 PR 30-JUN-1999; 99US-0141718P.
 XX
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Nunez G, Hu Y;
 XX
 DR WPI; 2001-112454/12.
 DR N-FSDB; AAA91116.
 XX
 PT Compositions for screening apoptosis pathway agonists and antagonists,
 PT useful for the treatment and diagnosis of cancer and autoimmune diseases,
 PT comprises new splice variants of wild-type Apaf-1 gene.
 XX
 PS Disclosure; Fig 26; 101pp; English.
 XX
 CC This sequence represents an Apaf-1 variant of the invention. The variants
 CC of the invention are all splice variants of a wild-type Apaf-1 gene (a
 CC mammalian homolog of CED-4 that participates in cytochrome c-dependent
 CC activation of caspase-3). The DNA is useful in screening assays for
 CC identifying apoptosis signalling pathway (antagonists, which are in turn
 CC useful as potential therapeutics and diagnostic or prognostic tools for
 CC diverse types of cancers, autoimmune diseases and hereditary diseases and
 CC for screening compounds that modulates the interaction of Apaf-1 with
 CC other members of the signalling pathway, i.e. their substrates or
 CC ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 CC cell-free assay system and the antibody generated to the translation
 CC product are used in immunoprecipitation assays to isolate new Apaf-1
 CC pathway constituents or their natural mutants

XX	Sequence 1205 AA;	
SQ	Query Match	
	Best Local Similarity 96.0%; Score 6352.5; DB 4; Length 1205;	
	Matches 1205; Conservative 0; Mismatches 0; Indels 43; Gaps 1;	
Qy	1 MDKARNCLLOHRALEKDKTSYVNDHMSDGLFTISEEEKVNEPTQOORAAMLIKMI 60	
Db	1 MDKARNCLLOHRALEKDKTSYVNDHMSDGLFTISEEEKVNEPTQOORAAMLIKMI 60	
Qy	61 LKXNDNSVSVFNALLHEGKDIKTSYVNDHMSDGLFTISEEEKVNEPTQOORAAMLIKMI 120	
Db	61 LKXNDNSVSVFNALLHEGKDIKTSYVNDHMSDGLFTISEEEKVNEPTQOORAAMLIKMI 120	
Qy	121 QRPVVFTRKLVNAIOOKSLKGBEGWTHGHAGCKSVLAFAVRDHSILLEGCPG 180	
Db	121 QRPVVFTRKLVNAIOOKSLKGBEGWTHGHAGCKSVLAFAVRDHSILLEGCPG 180	
Qy	181 GVHVSVGKQKSGLLMKLQNLCTRLDQDESQRPLNIEAKDRILMLRKHPRSL 240	
Db	181 GVHVSVGKQKSGLLMKLQNLCTRLDQDESQRPLNIEAKDRILMLRKHPRSL 240	
Qy	241 ILDDVWDSVNLKAFDSQCOILLTRDKSVTDSVMGPKYVPVSSSLGKEGELBILSLFVN 300	
Db	241 ILDDVWDSVNLKAFDSQCOILLTRDKSVTDSVMGPKYVPVSSSLGKEGELBILSLFVN 300	
Qy	301 MKKADLPQAHSHIIECKGSPVSVLIGALLRDPFNWMEYLLKOLNKPFRIRKSSSYD 360	
Db	301 MKKADLPQAHSHIIECKGSPVSVLIGALLRDPFNWMEYLLKOLNKPFRIRKSSSYD 360	
Qy	361 YEALDEAMSISVEMLRREDIKDYVTDLSILQKDVKPTKVLICILMDMETEEVEDILOEFVN 420	
Db	361 YEALDEAMSISVEMLRREDIKDYVTDLSILQKDVKPTKVLICILMDMETEEVEDILOEFVN 420	
Qy	421 KSLLFCDRNGKSFYRLHDLQVDFLTKNCSQDLHKKIITQFORVHQPHLSPQEDC 480	
Db	421 KSLLFCDRNGKSFYRLHDLQVDFLTKNCSQDLHKKIITQFORVHQPHLSPQEDC 480	
Qy	481 MYWYFLAYHMASAKMHELCALMFSLDWKAKTELVPAPHLIHEFVEYRHILDEKCAV 540	
Db	481 MYWYFLAYHMASAKMHELCALMFSLDWKAKTELVPAPHLIHEFVEYRHILDEKCAV 540	
Qy	541 SENFQEFSLNGHLLGRQFPFNIIVQLCEPTESEVYQAKLOAKQEVNDGMLYLEWINK 600	
Db	541 SENFQEFSLNGHLLGRQFPFNIIVQLCEPTESEVYQAKLOAKQEVNDGMLYLEWINK 600	
Qy	601 KNTNLRLVVRPHTDVYHACFSEDQRIASCADKTLQVFKATGSEKLEIKAHEDV 660	
Db	601 KNTNLRLVVRPHTDVYHACFSEDQRIASCADKTLQVFKATGSEKLEIKAHEDV 660	
Qy	661 LCCAFSTDDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGS 720	
Db	661 LCCAFSTDDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGS 720	
Qy	721 SDCLFLMDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKMDATSANERKS 780	
Db	721 SDCLFLMDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKMDATSANERKS 780	
Qy	781 INVKQFFLNLEDPQEDMEVIKCCSWSDGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 840	
Db	781 INVKQFFLNLEDPQEDMEVIKCCSWSDGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 840	
Qy	841 STIQYCDPSPQNHLLAVALSQYCVELWNTDSRSKVADCRGHLSSVHGVWFSFGSSFLTS 900	
Db	841 STIQYCDPSPQNHLLAVALSQYCVELWNTDSRSKVADCRGHLSSVHGVWFSFGSSFLTS 900	
Qy	901 SDDQTLRLWETKVCNSAVMLKQEVDFVQFQNEVWVLAVDHRLQLINGRTGQIDYLT 960	
Db	901 SDDQTLRLWETKVCNSAVMLKQEVDFVQFQNEVWVLAVDHRLQLINGRTGQIDYLT 960	
Qy	961 EAQVSCCLSPHLQYIAFGDENGAIETLELVNRRIFQSRFOHKKTVWHIQTADKTLIS 1020	

Db	918 EAQVSCCLSPHLQYIAFGDENGAIETLELVNRRIFQSRFOHKKTVWHIQTADKTLIS 977	
Qy	1021 SSDDAEIQVNNWQDKCIFIPLRGHQTVDKDFRLLKNSRLLSWSFDGTVKVWNIITGNKEKD 1080	
Db	978 SSDDAEIQVNNWQDKCIFIPLRGHQTVDKDFRLLKNSRLLSWSFDGTVKVWNIITGNKEKD 1037	
Qy	1081 FVCHQGTVLSCDISHDQATKFSSTADKTIKWSFDLLPLHELGRHNGCVRCSAFSDST 1140	
Db	1038 FVCHQGTVLSCDISHDQATKFSSTADKTIKWSFDLLPLHELGRHNGCVRCSAFSDST 1097	
Qy	1141 LIATGDDNGEIRIWNVSNGELLHLCAPISEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1200	
Db	1098 LIATGDDNGEIRIWNVSNGELLHLCAPISEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1157	
Qy	1201 WNVVVTGESSQTFYVTNGTNLKKIHYSPDKTVVTDNLGILYILQTL 1248	
Db	1158 WNVVVTGESSQTFYVTNGTNLKKIHYSPDKTVVTDNLGILYILQTL 1205	

Search completed: January 27, 2005, 18:36:24
Job time : 184 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 18:30:10 ; Search time 52 Seconds
(without alignments)
2309.202 Million cell updates/sec

Title: US-10-646-396-2
Perfect score: 6619
Sequence: 1 MDKARNCLLQHREALSKDI.....FKTVYVDNLGILVILQTL 1248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6282	94.9	1194	2 T03818	apoptotic proteina
2	732	11.1	1227	2 A21810	WD-40 repeat prote
3	728.5	11.0	1258	2 A12155	WD-repeat protei
4	691	10.4	1526	2 AC2239	WD-40 repeat prote
5	677.5	10.2	1711	2 AD1842	WD-40 repeat prote
6	672	10.2	1683	2 AF2071	WD-40 repeat prote
7	643.5	9.7	1356	2 T18521	beta transducin-li
8	626	9.5	1693	2 S76086	beta transducin-li
9	614.5	9.3	934	2 AG1889	WD-40 repeat prote
10	611.5	9.2	1551	2 AB2410	WD-repeat protei
11	610	9.2	1708	2 AE1866	WD-40 repeat prote
12	600.5	9.1	1189	2 A12493	WD-repeat protei
13	580.5	8.8	1189	2 AH2154	WD-repeat protei
14	561.5	8.5	1747	2 AC1842	WD-40 repeat prote
15	464.5	7.0	1329	2 A21901	WD-repeat containi
16	441.5	6.7	1101	2 T26919	hypothetical prote
17	423	6.4	2629	2 T32735	telomerase-associ
18	399.5	6.0	2629	2 T30987	telomerase-associ
19	396	6.0	1049	2 T42045	beta transducin-li
20	394	6.0	777	2 T41075	hypothetical WD-re
21	391.5	5.9	1191	2 S76414	beta transducin-li
22	385	5.8	677	2 A21861	serine/threonine k
23	370	5.6	317	2 T46032	WD-40 repeat regul
24	369	5.6	786	2 AG2375	WD-40 repeat prote
25	365	5.5	676	2 AH2195	hypothetical prote
26	355	5.4	876	2 T51507	WD40-repeat protei
27	338.5	5.1	589	2 AG2400	WD-repeat protei
28	326	4.9	559	2 AB2202	hypothetical prote
29	322.5	4.9	515	2 S19487	hypothetical prote

ALIGNMENTS

RESULT 1

T03818

apoptotic proteinase-activating factor 1 - human

N:Alternate names: APAF-1 protein; protein KIAA0413

C:Species: Homo sapiens (man)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: T03818; T00053

R:Zou, H.; Henzel, W.J.; Liu, X.S.; Lutschg, A.; Wang, X.D.

Cell 90, 405-413, 1997

A:Title: Apaf-1, a human protein homologous to C. elegans CED-4, participates in cytochr

A:Reference number: Z15099; MUID:97410306; PMID:9267021

A:Accession: T03818

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-1194 <ZOU>

A:Cross-references: UNIPROT:O14727; EMBL:AF013263; NID:g2330014; PIDN:AA51678.1; PID:g2

A:Experimental source: cell line HeLa S3

R:Ishikawa, K.; Nagase, T.; Nakajima, D.; Seki, N.; Ohira, M.; Miyajima, N.; Tanaka, A.,

submitted to the EMBL Data Library, October 1997

A:Description: Prediction of the coding sequences of unidentified human genes. VIII. The

A:Reference number: Z14080

A:Accession: T00053

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 374-812, 'FDIHTSGLLGELIHTGHSTIQYCDPSQNHVAVVALSYQYCVEL', 813-1186, 'I', 1188-1194

A:Cross-references: EMBL:AB007873; NID:d1179749; PIDN:BAA24843.1; PID:d1025765

A:Experimental source: brain

C:Genetics:

A:Gene: Apaf-1

A:Note: KIAA0413

C:Function:

A:Description: binds and hydrolyzes ATP or dATP to ADP or dADP, respectively; participat

Query Match 94.9%; Score 6282; DB 2; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

Qy	1	MDAKARNCLLQHREALKDIKTSYIMDHMSDGLTITSEEEKVRNEPTQQQRAAMLIKMI	60
Db	1	MDAKARNCLLQHREALKDIKTSYIMDHMSDGLTITSEEEKVRNEPTQQQRAAMLIKMI	60
Qy	61	LKNDNDVSVPYNALLHEGYKDLAALLHDGIPVYSSSSGKDSVSGITTSYVTVLCEGGVP	120
Db	61	LKNDNDVSVPYNALLHEGYKDLAALLHDGIPVYSSSSGKDSVSGITTSYVTVLCEGGVP	109
Qy	121	QRPVVFVTRKLVNAIQKLSKLGEPGWVTHGMAGCGKSVLAEEAVRDSHSLLEGCPFG	180
Db	110	QRPVVFVTRKLVNAIQKLSKLGEPGWVTHGMAGCGKSVLAEEAVRDSHSLLEGCPFG	169
Qy	181	GVHWSVGKQDKSLIMKLNLCITRLDQDSFSQRLPLNIEAKDRILMLRKHPRSL	240
Db	170	GVHWSVGKQDKSLIMKLNLCITRLDQDSFSQRLPLNIEAKDRILMLRKHPRSL	229

241 ILDDVMDSWLKAQDSQOILLTRDKSVTDSVMGPKYVVPVSSSLGKGLBILSLFVN 300
 Db ILDDVMDSWLKAQDSQOILLTRDKSVTDSVMGPKYVVPVSSSLGKGLBILSLFVN 289
 301 MKKADLPQAHSHIIECKGSPVLSLIGALLRDFNREYILKQNKQPKRIRKSSSYD 360
 Db MKKADLPQAHSHIIECKGSPVLSLIGALLRDFNREYILKQNKQPKRIRKSSSYD 349
 361 YEALDEAMSIIVEMREDIKDYVTDLSILQKQVPTKVLGILMDMETEEVEDILOEFVN 420
 Db YEALDEAMSIIVEMREDIKDYVTDLSILQKQVPTKVLGILMDMETEEVEDILOEFVN 409
 421 KSLIFCDRNGKSFYRLHDLQVDFLEKNCQDLHKKIITQFORVHOPHTLSPOQEDC 480
 Db KSLIFCDRNGKSFYRLHDLQVDFLEKNCQDLHKKIITQFORVHOPHTLSPOQEDC 469
 481 MYWYNFLAYHWSAKMKECALMFSLDWKATLGVGPAHLIHEFVEYRHLDEKCAV 540
 Db MYWYNFLAYHWSAKMKECALMFSLDWKATLGVGPAHLIHEFVEYRHLDEKCAV 529
 541 SENFOEPLSINGHLIGROPENIVOLGCEPETSEVYQAKLOAKQEVNDGMLYLEWINK 600
 Db SENFOEPLSINGHLIGROPENIVOLGCEPETSEVYQAKLOAKQEVNDGMLYLEWINK 589
 601 KMITNLRLVPRDHTDVAHYHACSEDEGRIASCADKTLQVKAETGEKLEIKAHDEVE 660
 Db KMITNLRLVPRDHTDVAHYHACSEDEGRIASCADKTLQVKAETGEKLEIKAHDEVE 649
 661 LCCAFSTDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
 Db LCCAFSTDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 709
 721 SFCFLKLDLQKQCRNMTGHTNSVNHCRFPDPDKLLASCADGTLKLDATSANERKS 780
 Db SFCFLKLDLQKQCRNMTGHTNSVNHCRFPDPDKLLASCADGTLKLDATSANERKS 769
 781 INVKQFFNLNEDPQEDMEVIVKCCSWADGARIWAAKNKIFLEDIHTSGLLGEIHTGHH 840
 Db INVKQFFNLNEDPQEDMEVIVKCCSWADGARIWAAKNKIFLEDIHTSGLLGEIHTGHH 811
 841 STIQYCDSPQWHLAVALSQYCVELWNTDSRSKADCRHLSVHWGVPSPDGSFLTS 900
 Db STIQYCDSPQWHLAVALSQYCVELWNTDSRSKADCRHLSVHWGVPSPDGSFLTS 846
 901 SDDQIRLWETKVKCKSAVWLKQEDVVDVVOENWVWLAVDHRLQOLINGRTQIDYLT 960
 Db SDDQIRLWETKVKCKSAVWLKQEDVVDVVOENWVWLAVDHRLQOLINGRTQIDYLT 906
 961 EAQVSCCLSPHLQYIAFGDENGAIETELVNNRIFQSFQHKKTVMHIQFTADEKTLIS 1020
 Db EAQVSCCLSPHLQYIAFGDENGAIETELVNNRIFQSFQHKKTVMHIQFTADEKTLIS 966
 1021 SDDDAEIQWNNQDLCIFLAGHOETVKOPRLLKNSRLLSRFSFGTVKWNIIITGNKEVD 1080
 Db SDDDAEIQWNNQDLCIFLAGHOETVKOPRLLKNSRLLSRFSFGTVKWNIIITGNKEVD 1026
 1081 FVCHQSTVLSGDISHDATKFSSTSAKTAKIWSFDLLPLHELGHNGCVRCGAFSDVT 1140
 Db FVCHQSTVLSGDISHDATKFSSTSAKTAKIWSFDLLPLHELGHNGCVRCGAFSDVT 1086
 1141 LLATGDDNGEIRIWNVNGELLHLCAPLSEGAATHGGVWTDLCFSPDGKWLISAGGYIK 1200
 Db LLATGDDNGEIRIWNVNGELLHLCAPLSEGAATHGGVWTDLCFSPDGKWLISAGGYIK 1146
 1201 WNNVVTGSSQFTYNGTNLKKIHVSFDFKTYVVDNGLIYLQTL 1248
 Db WNNVVTGSSQFTYNGTNLKKIHVSFDFKTYVVDNGLIYLQTL 1194
 1147 WNNVVTGSSQFTYNGTNLKKIHVSFDFKTYVVDNGLIYLQTL 1194

RESULT 2
 AE1810
 WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C;Accession: AE1810
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AE1810
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1227 <KUR>
 A;Cross-references: UNIPROT:Q8Z0R1; GB:BA000019; PIDN:BA077553.1; PID:gl7135007; GSPDB:G
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: alr0029

Query Match 11.1%; Score 732; DB 2; Length 1227;
 Best Local Similarity 22.2%; Pred. No. 1.3e-35;
 Matches 295; Conservative 209; Mismatches 438; Indels 384; Gaps 50;

QY 77 HEGYKDLAALLHOGIPVSSSSGKDSVSGITSYVTVLCEGVPQRPVFWTRKKLVAI 136
 Db 45 HSG-AETAAKLDISQAVRKLGE-----SY-RKLGIQ-----KGNKINGL 85
 QY 137 QOKLSK-----LKGSPGW-----VTIHG 154
 Db 86 ROKLYEQVLSYQHPVSSSEDGCEAVDVAGFRGRKEPILLEQWIEGNSNRCRLVAVLG 145
 QY 155 MAGCKSVLAEEAVR-----DHSLLGCGFGGWHVSVGKQKSG-LLMKLONLCTRLD 207
 Db 146 MGGTGKTVLAAMGAKKQVKEFDYLI-----WLSRLNAPSLGDLITLQLLRFLANDN 195
 QY 208 QDESFSQRLPANTEAEKADRLRLMLRKHPRLSLILDDVW-----DSW----- 249
 Db 196 END-----LTDTDNNKIVRLDVLKRCRLIILDNVESVLSRSGEKQEWAGDYQGY 248
 QY 250 -----VLKAFDSQCOQLLTTRDKSVTDSVMGPKYVVPVSSSLGKGLBILSL-FV 299
 Db 249 ENYGFLFKVAEASHESC-LLLTSTREK-----PKEVAAL-----GKNLPVKVQLSSL 296
 QY 300 NKKKAD-----LPEQAHSHIIECKGSPVLSLIGALLRD-FPNRWEYILKQ----- 344
 Db 297 NLAEAREILLDKGYCTDEQLDELVRYSNGNPLAKIVATTVVELFSNNISEFLTQIHOE 356
 QY 345 -----LQNKQFKR-----IRKSSSYDEALDEAMSIIVEMREDIKOYVTDLSILQXD 392
 Db 357 SAVYGDITLLKQOFKRLSELEKVMYSLGANREYVS-----FRELKDDMLTTES----- 406
 QY 393 VKVPTKVLCLMDMETEEVEDILOEFV-----NKSLLFCDRNGKSFYRLHDLQV 442
 Db 407 ---PIKVM-----EALLESLLRSLIEKASPTLIEKASSTQGEKEAESKFGLESVVM 455
 QY 443 DFLTEKNCQDLHKKIITQFORVHOPHTLSPOQEDCMYWNFL-AYHMASAKMKEKLC 501
 Db 456 EYTA-----KEIENSLEESQTKL-----DFINTYPLMKAR----- 488
 QY 502 ALMPSLDWIKAKTELVPAPHLIHEFVEYRHLDEKDCAVSENQOEFLSLNG-----HL-- 554
 Db 489 ----SLDYIROIQE-----RLILEPVK-----QKLNIFGTLELEHLRR 523
 QY 555 ----IGROPPP-----NIV-----QLGLCE-PETSEV-----YQOAKL 582
 Db 524 MLGTLOKEPLPKGYAGNLIINLLRQLDKIPDESPIJLSGRDFSGLTITWQAYFKEVKL 583
 QY 583 Q-----AKQEVNDGMLYLEWINK-----KNITNLRLVVRP 613
 Db 584 KETIFANSDL-TGSVFTETMNSVSVKFSFGPKGYPATGLMNGEIRLWQTSNDKQLRIYKG 642
 QY 614 HTDAVTHACFSEDEGRIASCADKTLQVKAETGEKLEIKAHDEVEYLCCAFSTDRFTA 673
 Db 643 HTAWWFAFSPDSRMLASGSDSTIKLMDVHTGBCLTLSKNTNKVYSVAFSPDGRILA 702

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 18:40:29 ; Search time 172 Seconds
(without alignments)
2621.449 Million cell updates/sec

Title: US-10-646-396-2
Perfect score: 6619
Sequence: 1 MDAKARNCLLOHREALEKDI.....PKTVTVVDNLGILVLOTLE 1248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6619	100.0	1248	17	US-10-482-952-16
2	6352.5	96.0	1205	9	US-09-876-667-16
3	6282	94.9	1194	9	US-09-876-667-2
4	6282	94.9	1194	14	US-10-141-618-10
5	6003.5	90.7	1199	15	US-10-112-944-273
6	5933.5	89.6	1526	15	US-10-112-944-272
7	1589	24.0	308	13	US-10-014-269-31
8	1589	24.0	308	13	US-10-002-974-31
9	1589	24.0	308	14	US-10-314-506-31
10	748	11.3	142	14	US-10-001-254-29
11	711	10.7	529	14	US-10-369-493-12913
12	690.5	10.4	610	14	US-10-369-493-20185
13	688	10.4	1005	14	US-10-369-493-20242

14	674	10.2	1140	14	US-10-369-493-19053	Sequence 19053, A
15	672	10.2	1136	14	US-10-369-493-19046	Sequence 19046, A
16	652	9.9	579	14	US-10-369-493-18893	Sequence 18893, A
17	643.5	9.7	1356	13	US-10-077-111-10	Sequence 10, Appl
18	632.5	9.6	1118	14	US-10-369-493-18845	Sequence 18845, A
19	606.5	9.2	1155	14	US-10-369-493-19869	Sequence 19869, A
20	604.5	9.1	1117	14	US-10-369-493-18965	Sequence 18965, A
21	577	8.7	608	14	US-10-369-493-20224	Sequence 20224, A
22	567	8.6	613	14	US-10-369-493-11181	Sequence 1181, A
23	555.5	8.4	606	14	US-10-369-493-18907	Sequence 18907, A
24	552	8.3	625	14	US-10-369-493-20066	Sequence 20066, A
25	526.5	8.0	656	14	US-10-369-493-20247	Sequence 20247, A
26	526.5	8.0	679	14	US-10-369-493-20059	Sequence 20059, A
27	522.5	7.9	478	14	US-10-369-493-18905	Sequence 18905, A
28	496.5	7.5	574	14	US-10-369-493-20058	Sequence 20058, A
29	483	7.3	584	14	US-10-369-493-19944	Sequence 19944, A
30	476	7.2	93	13	US-10-014-269-28	Sequence 28, Appl
31	476	7.2	93	13	US-10-002-974-28	Sequence 28, Appl
32	476	7.2	93	14	US-10-314-506-28	Sequence 28, Appl
33	461	7.0	90	14	US-10-156-733-4	Sequence 4, Appl
34	455	6.9	89	9	US-09-841-739-16	Sequence 16, Appl
35	455	6.9	89	9	US-09-931-071-8	Sequence 8, Appl
36	444.5	6.7	429	14	US-10-449-315-16	Sequence 16, Appl
37	444.5	6.7	429	14	US-10-369-493-11348	Sequence 11348, A
38	430	6.5	1102	14	US-10-369-493-19941	Sequence 19941, A
39	423	6.4	2629	14	US-10-295-681-54	Sequence 54, Appl
40	423	6.4	2629	14	US-10-295-681-55	Sequence 55, Appl
41	423	6.4	2629	14	US-10-295-681-61	Sequence 61, Appl
42	410	6.2	2630	15	US-10-334-143-41	Sequence 41, Appl
43	368.5	5.6	559	14	US-10-369-493-20028	Sequence 20028, A
44	366	5.5	891	16	US-10-437-963-153705	Sequence 153705, A
45	358	5.4	280	14	US-10-369-493-19899	Sequence 19899, A

ALIGNMENTS

RESULT 1
US-10-482-952-16
; Sequence 16, Application US/10482952
; Publication No. US20040254136A1
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Columbia University in the City of New York, et al.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES AND RELATED METHODS FOR REGULATING CEI
; FILE REFERENCE: 0575/64078-A-PCT/JPW/ANX
; CURRENT APPLICATION NUMBER: US/10/482.952
; CURRENT FILING DATE: 2004-01-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 16
; LENGTH: 1248
; TYPE: PRT
; ORGANISM: human
US-10-482-952-16

Query Match 100.0%; Score 6619; DB 17; Length 1248;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MDAKARNCLLOHREALEKDIKTSYIMDHIMISDGFLLTSEEKVRNEPTQQOQAAMLIKMI	60
Db	1	MDAKARNCLLOHREALEKDIKTSYIMDHIMISDGFLLTSEEKVRNEPTQQOQAAMLIKMI	60
Qy	61	LKNDNDVSVPYNALLHEGYKDLAALLHDGIPVYSSSSGKDSVSGITSYVTVLCEGVP	120
Db	61	LKNDNDVSVPYNALLHEGYKDLAALLHDGIPVYSSSSGKDSVSGITSYVTVLCEGVP	120
Qy	121	QRPVVFVTRKKLVNAIQKLSKLGEGPWVTHGMAGCKSVLAEEAVRDSHSLLEGCFPG	180
Db	121	QRPVVFVTRKKLVNAIQKLSKLGEGPWVTHGMAGCKSVLAEEAVRDSHSLLEGCFPG	180
Qy	181	GVHWSVGCKQKSGILLMKLQNLCTRLDODESFSORPLNIEAKDRLILMLRKHPRSL	240

```

Db 181 GVHWVSGVQDKSGLLMKLQNLCTRLDQDESPQRPLNIEBAKDRILMLRKHPRSL 240
Qy 241 ILDDVWDSWLVKAFDSQCIILLTRDKSVTDSVMGPKYVVPVSSSLGKGLKLEILSLFN 300
Db 241 ILDDVWDSWLVKAFDSQCIILLTRDKSVTDSVMGPKYVVPVSSSLGKGLKLEILSLFN 300
Qy 301 MKKADLPEQAHSIKECKGSPVLSLIGALLRDPFNNRMEYLLKQKQKFRIRKSSSYD 360
Db 301 MKKADLPEQAHSIKECKGSPVLSLIGALLRDPFNNRMEYLLKQKQKFRIRKSSSYD 360
Qy 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKOVKVPKVLICILMDMETEEVEDILOEFVN 420
Db 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKOVKVPKVLICILMDMETEEVEDILOEFVN 420
Qy 421 KSLIFCDNRNGKFRYLLHDLQVDFLTENKCSQDLHKKIITQFORVHOPHTLSPQDEDC 480
Db 421 KSLIFCDNRNGKFRYLLHDLQVDFLTENKCSQDLHKKIITQFORVHOPHTLSPQDEDC 480
Qy 481 MYWYNFLAYHMASAKWHELCALMFSLDWIKAKTELVGPAHLIHEFVEYRHILDEKDCAV 540
Db 481 MYWYNFLAYHMASAKWHELCALMFSLDWIKAKTELVGPAHLIHEFVEYRHILDEKDCAV 540
Qy 541 SENFQEFSLNGHLGROPENIVOLGLCEPETSEVYQOAKLOKQVNDGMLYLEWINK 600
Db 541 SENFQEFSLNGHLGROPENIVOLGLCEPETSEVYQOAKLOKQVNDGMLYLEWINK 600
Qy 601 KNTNLSRLVVRPHTDAVTHACFSEDQORTASCADKTLQVFKARTGKLEIKAHEDV 660
Db 601 KNTNLSRLVVRPHTDAVTHACFSEDQORTASCADKTLQVFKARTGKLEIKAHEDV 660
Qy 661 LCCAFSTDDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
Db 661 LCCAFSTDDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
Qy 721 SDCFLKMLDNLQKCRNTMFHTNNSVNHCRFSPDDKLLASCADGTLKMDATSANERKS 780
Db 721 SDCFLKMLDNLQKCRNTMFHTNNSVNHCRFSPDDKLLASCADGTLKMDATSANERKS 780
Qy 781 INVQKPFNLDPQEDMEVIVKCCSWADGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 840
Db 781 INVQKPFNLDPQEDMEVIVKCCSWADGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 840
Qy 841 STIQYCDFSPQNLAVVALSYQVELWNTDSRSKVADCRGHLVWVGFSPDGSSFLTS 900
Db 841 STIQYCDFSPQNLAVVALSYQVELWNTDSRSKVADCRGHLVWVGFSPDGSSFLTS 900
Qy 901 SDDQIRLWETKVKVCKNSAVMLKQEDVDFQENVMVLAVDHRRLOLQINGRTGQIDYLT 960
Db 901 SDDQIRLWETKVKVCKNSAVMLKQEDVDFQENVMVLAVDHRRLOLQINGRTGQIDYLT 960
Qy 961 EAQVSCCCLSPHLQYIAGDENGAIETLELVNRRIFQSRFOHKKTVWHIOPTADEKTLIS 1020
Db 961 EAQVSCCCLSPHLQYIAGDENGAIETLELVNRRIFQSRFOHKKTVWHIOPTADEKTLIS 1020
Qy 1021 SSDDAEIOVNNWOLDKCIFLRGHQTVDKOPRLLKNSRLLSWSFDGTWKVNNIITGNKED 1080
Db 1021 SSDDAEIOVNNWOLDKCIFLRGHQTVDKOPRLLKNSRLLSWSFDGTWKVNNIITGNKED 1080
Qy 1081 FVCHQGTVLSCDISHDATKSSISADTKATKISFDLLIPLHELGHNGCVRCASFVDS 1140
Db 1081 FVCHQGTVLSCDISHDATKSSISADTKATKISFDLLIPLHELGHNGCVRCASFVDS 1140
Qy 1141 LLAATDNGEIRINWVNSGELLHLHCAPLSEGAATHGQVWTDLCFSPDGKMLISAGYIK 1200
Db 1141 LLAATDNGEIRINWVNSGELLHLHCAPLSEGAATHGQVWTDLCFSPDGKMLISAGYIK 1200
Qy 1201 WNNVVTGESSQFTYNGTGNLKKIHVSPDKTYVTVDNLGILYILOTLE 1248
Db 1201 WNNVVTGESSQFTYNGTGNLKKIHVSPDKTYVTVDNLGILYILOTLE 1248

```

RESULT 2

US-09-876-667-16

```

; Sequence 16, Application US/09876667
; Patent No. US20020107370A1
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. US20020107370A1west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/876,667
; FILING DATE: 07-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/435,115
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; -TYPE: amino acid
; -LENGTH: 1205 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
;
US-09-876-667-16
;
Query Match 96.0%; Score 6352.5; DB 9; Length 1205;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1205; Conservative 0; Mismatches 0; Indels 43; Gaps 1;
Qy 1 MDKARNCLLQHREALEKDKITSYIMDHMISDGFLLTISEBEKVRNEPTQOORAAMLKMI 60
Db 1 MDKARNCLLQHREALEKDKITSYIMDHMISDGFLLTISEBEKVRNEPTQOORAAMLKMI 60
Qy 61 LKKONDSVSVFYNALHREGVKDLAALLHDGIPVSVSSSGKDSVSGITSYVRTVLCGGVP 120
Db 61 LKKONDSVSVFYNALHREGVKDLAALLHDGIPVSVSSSGKDSVSGITSYVRTVLCGGVP 120
Qy 121 QRPVVFVTRKLVNATOOKSLKGERGVNTHMGACGKSVLAARVDRHSLLEGCPG 180
Db 121 QRPVVFVTRKLVNATOOKSLKGERGVNTHMGACGKSVLAARVDRHSLLEGCPG 180
Qy 181 GVHWVSGVKQDKSGLLMKLQNLCTRLDQDESPQRPLNIEBAKDRILMLRKHPRSL 240
Db 181 GVHWVSGVKQDKSGLLMKLQNLCTRLDQDESPQRPLNIEBAKDRILMLRKHPRSL 240
Qy 241 ILDDVWDSWLVKAFDSQCIILLTRDKSVTDSVMGPKYVVPVSSSLGKGLKLEILSLFN 300
Db 241 ILDDVWDSWLVKAFDSQCIILLTRDKSVTDSVMGPKYVVPVSSSLGKGLKLEILSLFN 300
Qy 301 MKKADLPEQAHSIKECKGSPVLSLIGALLRDPFNNRMEYLLKQKQKFRIRKSSSYD 360
Db 301 MKKADLPEQAHSIKECKGSPVLSLIGALLRDPFNNRMEYLLKQKQKFRIRKSSSYD 360

```

```

Qy 361 YEALDEAMSISVEMLRREDIKDYITDLSILQKDVKVPKVLKILMDMETEEVEDILOQFVN 420
Db 361 YEALDEAMSISVEMLRREDIKDYITDLSILQKDVKVPKVLKILMDMETEEVEDILOQFVN 420
Qy 421 KSLLFCDNRNGSKFRYLLHDLQVDFLTERKNSQLQDLHKKIITQFQRYHQPHTLSPDQEDC 480
Db 421 KSLLFCDNRNGSKFRYLLHDLQVDFLTERKNSQLQDLHKKIITQFQRYHQPHTLSPDQEDC 480
Qy 481 MYWNLFLAYMASAKHKLALCALMFSLDWIRAKTELVGPAHLIHEFVYRHLDEKDCAV 540
Db 481 MYWNLFLAYMASAKHKLALCALMFSLDWIRAKTELVGPAHLIHEFVYRHLDEKDCAV 540
Qy 541 SENQOEFUSLNGHLIGROFFNIVOLGICEPSETSEVYOQAKLOAQKQEVNDGMLYLEWINK 600
Db 541 SENQOEFUSLNGHLIGROFFNIVOLGICEPSETSEVYOQAKLOAQKQEVNDGMLYLEWINK 600
Qy 601 KNIITNLSRLVVRPHTDAVYHACFSDGRIASCAGADKTLQVFKAGTGEKLEIKAHEDV 660
Db 601 KNIITNLSRLVVRPHTDAVYHACFSDGRIASCAGADKTLQVFKAGTGEKLEIKAHEDV 660
Qy 661 LCCAFSTDDRIATCSVDKVKIWNMTGELVHYDEHSEOVNCHFTNSHHLLLATGS 720
Db 661 LCCAFSTDDRIATCSVDKVKIWNMTGELVHYDEHSEOVNCHFTNSHHLLLATGS 720
Qy 721 SDCFLKLDLWLNQECRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 780
Db 721 SDCFLKLDLWLNQECRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 780
Qy 781 INVQOFFLNLDEPQEDMEVIVKCCSWSDAGARIWVAANKIFLFDIHTSGLLGIHTGHH 840
Db 781 INVQOFFLNLDEPQEDMEVIVKCCSWSDAGARIWVAANKIFLFDIHTSGLLGIHTGHH 840
Qy 841 STIOYCDPSPQNLAVVALSYCVELWNTDGRSKVADCRGHLVWVHGMFSPDGSSFLT 900
Db 841 STIOYCDPSPQNLAVVALSYCVELWNTDGRSKVADCRGHLVWVHGMFSPDGSSFLT 900
Qy 882 INVQOFFLNLDEPQEDMEVIVKCCSWSDAGARIWVAANKIFLFDIHTSGLLGIHTGHH 882
Db 882 INVQOFFLNLDEPQEDMEVIVKCCSWSDAGARIWVAANKIFLFDIHTSGLLGIHTGHH 882
Qy 901 SDDQIRLWETKVKCKNSAVMLKQEVVVFQENVMVLAVDHIRRLQLINGRTQOIDL 960
Db 901 SDDQIRLWETKVKCKNSAVMLKQEVVVFQENVMVLAVDHIRRLQLINGRTQOIDL 960
Qy 961 EAQVSCCLSLHQLYIAFGDENGAIIELELVNRIQSRFOHKTVMWHIQTADEKTLIS 1020
Db 961 EAQVSCCLSLHQLYIAFGDENGAIIELELVNRIQSRFOHKTVMWHIQTADEKTLIS 1020
Qy 1021 SSDDAEIOVWNLQDKCIFLGHQETVKDFRLLKNSRLLSVDFGTVKVWNIITGNKEK 1080
Db 1021 SSDDAEIOVWNLQDKCIFLGHQETVKDFRLLKNSRLLSVDFGTVKVWNIITGNKEK 1080
Qy 1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLPLHELGHNGCVRCSPVSDST 1140
Db 1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLPLHELGHNGCVRCSPVSDST 1140
Qy 1141 LLATGDDNGETIRIWNVSGELHLICAPLSEGAATHGQWVTDLCSPDGKMLISAGGYIK 1200
Db 1141 LLATGDDNGETIRIWNVSGELHLICAPLSEGAATHGQWVTDLCSPDGKMLISAGGYIK 1200
Qy 1201 WNVVTGSSOTFYTNGNLKIHVSDFKTYVTVDNLGILYILOTLE 1248
Db 1201 WNVVTGSSOTFYTNGNLKIHVSDFKTYVTVDNLGILYILOTLE 1248
Qy 1248 WNVVTGSSOTFYTNGNLKIHVSDFKTYVTVDNLGILYILOTLE 1248
Db 1248 WNVVTGSSOTFYTNGNLKIHVSDFKTYVTVDNLGILYILOTLE 1248

```

RESULT 3

```

US-09-876-667-2
; Sequence 2, Application US/09876667
; Patent No. US20020107370A1
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAP-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. US20020107370A1west Center, 90 South Seventh St

```

```

; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/876,667
; FILING DATE: 07-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/435,115
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-876-667-2

```

```

Query Match 94.9%; Score 6282; DB 9; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

```

```

Qy 1 MDKARNCCLLOHRALEKDKITSYIMDMHISDGLTITSEEEKVNEPTQOQRAAMLIKMI 60
Db 1 MDKARNCCLLOHRALEKDKITSYIMDMHISDGLTITSEEEKVNEPTQOQRAAMLIKMI 60
Qy 61 LKNDNSVSVFNALLHEGYKDLAALLHDGIPVYVSSSSGKDSVSGITSYVTVLCEGVP 120
Db 61 LKNDNSVSVFNALLHEGYKDLAALLHDGIPVYVSSSSGKDSVSGITSYVTVLCEGVP 120
Qy 121 QRPVVFVTRKLVNAIOQKLSKLKGEFGWVTHGMAGCGKSVLAEEAVRDHSLLEGCPFG 180
Db 121 QRPVVFVTRKLVNAIOQKLSKLKGEFGWVTHGMAGCGKSVLAEEAVRDHSLLEGCPFG 180
Qy 181 GVHWSVVKODKSGLLMKLQNLCTRLDQDSFSQRLPLNTEAKDRILMLRKHPSLL 240
Db 181 GVHWSVVKODKSGLLMKLQNLCTRLDQDSFSQRLPLNTEAKDRILMLRKHPSLL 240
Qy 241 ILDDVDMSWLKAFDSCQIILLTRDKSVTDSVNGPKYVVPVSSSLGKEGLEILSLFVN 300
Db 241 ILDDVDMSWLKAFDSCQIILLTRDKSVTDSVNGPKYVVPVSSSLGKEGLEILSLFVN 300
Qy 301 MKKADLPQAHSIIECKGSPVLSLIGALLRDPFNRWEYLLKQLQNKQPKRIRKSSSYD 360
Db 301 MKKADLPQAHSIIECKGSPVLSLIGALLRDPFNRWEYLLKQLQNKQPKRIRKSSSYD 360
Qy 361 YEALDEAMSISVEMLRREDIKDYITDLSILQKDVKVPKVLKILMDMETEEVEDILOQFVN 420
Db 361 YEALDEAMSISVEMLRREDIKDYITDLSILQKDVKVPKVLKILMDMETEEVEDILOQFVN 420
Qy 421 KSLLFCDNRNGSKFRYLLHDLQVDFLTERKNSQLQDLHKKIITQFQRYHQPHTLSPDQEDC 480
Db 421 KSLLFCDNRNGSKFRYLLHDLQVDFLTERKNSQLQDLHKKIITQFQRYHQPHTLSPDQEDC 480
Qy 480 KSLLFCDNRNGSKFRYLLHDLQVDFLTERKNSQLQDLHKKIITQFQRYHQPHTLSPDQEDC 480
Db 480 KSLLFCDNRNGSKFRYLLHDLQVDFLTERKNSQLQDLHKKIITQFQRYHQPHTLSPDQEDC 480

```

Qy	481	MYWTFLAYHMAKQKH	ELCALMFSLDNT	IKATELVGPAHL	IHFEVYVYHILDEKCAV	540																									
Db	470	MYWTFLAYHMAKQKH	ELCALMFSLDNT	IKATELVGPAHL	IHFEVYVYHILDEKCAV	529																									
Qy	541	SENFOEFLSUNGHL	GRQFPFN	IOLGLCBP	SEFSEVYQOAKLQAKQBVNDGMLYLEWLNK	600																									
Db	530	SENFOEFLSUNGHL	GRQFPFN	IOLGLCBP	SEFSEVYQOAKLQAKQBVNDGMLYLEWLNK	589																									
Qy	601	KNIITNLSRLVVR	PHTDVA	VYHACFSD	GGRIASCGADKTLOVFKAETGKELLEIKAHEDV	660																									
Db	590	KNIITNLSRLVVR	PHTDVA	VYHACFSD	GGRIASCGADKTLOVFKAETGKELLEIKAHEDV	649																									
Qy	661	LCCAPSTDD	RFIATCS	VDKVKIWN	SMTGELVHTYDHESEQVNCCHPTNSSHHLLLATGS	720																									
Db	650	LCCAPSTDD	RFIATCS	VDKVKIWN	SMTGELVHTYDHESEQVNCCHPTNSSHHLLLATGS	709																									
Qy	721	SDCFLKLWD	LNQKECRNT	MFGHTNS	VNHCFCFSPDDKLLASC	SADGTLKLWDATSANERKS	780																								
Db	710	SDCFLKLWD	LNQKECRNT	MFGHTNS	VNHCFCFSPDDKLLASC	SADGTLKLWDATSANERKS	769																								
Qy	781	INVQFFINLE	POQEDME	VYKCCSN	SADGARIMVAAKNKIFLEPDIHTSGLLGEIHTGH	840																									
Db	770	INVQFFINLE	POQEDME	VYKCCSN	SADGARIMVAAKNKIF-----	811																									
Qy	841	STIQCYD	FSQNH	LAVALS	OYCELVNMTD	DSRSKVADCRGHS	LWVHGWVWFSFGDSSFLT	900																							
Db	812	-----	-----	-----	-----	-----	-----	846																							
Qy	901	SDDOTIR	LEWTKVK	CKNSAV	MLKEQVDV	FQENEVWVLAVD	HIRRLQINGRTGQIDYLT	960																							
Db	847	SDDOTIR	LEWTKVK	CKNSAV	MLKEQVDV	FQENEVWVLAVD	HIRRLQINGRTGQIDYLT	906																							
Qy	961	BAQVSCC	ISPHLQ	IAFG	DENGAIE	ILELVNNR	IFQSRFOHKKTVMH	IQFTADSEKTLIS	1020																						
Db	907	BAQVSCC	ISPHLQ	IAFG	DENGAIE	ILELVNNR	IFQSRFOHKKTVMH	IQFTADSEKTLIS	966																						
Qy	1021	SSDDAE	IQVNNQ	LQDKC	IFLRGHQ	ETVKOP	FRLLKNS	RLLSWSFGD	TVKVMNII	1080																					
Db	967	SSDDAE	IQVNNQ	LQDKC	IFLRGHQ	ETVKOP	FRLLKNS	RLLSWSFGD	TVKVMNII	1026																					
Qy	1081	FVCHQGT	VLSC	DIS	HATKFS	ST	SADTKI	WFS	FDLLPL	HELGRHNG	CVRC	SAF	SVDST	1140																	
Db	1027	FVCHQGT	VLSC	DIS	HATKFS	ST	SADTKI	WFS	FDLLPL	HELGRHNG	CVRC	SAF	SVDST	1086																	
Qy	1141	LLATGDD	NGEIR	IR	VNWS	NGELLH	CAP	LS	BEGA	ATHG	QW	TD	LC	FSP	DG	ML	I	SAGGYIK	1200												
Db	1087	LLATGDD	NGEIR	IR	VNWS	NGELLH	CAP	LS	BEGA	ATHG	QW	TD	LC	FSP	DG	ML	I	SAGGYIK	1146												
Qy	1201	WNVVT	G	SSO	F	Y	T	N	G	L	I	H	V	S	P	D	K	T	V	V	D	N	I	G	I	L	O	T	L	E	1248
Db	1147	WNVVT	G	SSO	F	Y	T	N	G	L	I	H	V	S	P	D	K	T	V	V	D	N	I	G	I	L	O	T	L	E	1194

RESULT 4

```

RESULT= 4
US/10/141-618-10
, Sequence 10, Application US/10141618
, Publication No. US20030165897A1
, GENERAL INFORMATION:
, APPLICANT: Reed, John C.
, TITLE OF INVENTION: Methods For Determining
, TITLE OF INVENTION: For Cancer Patients
, FILE REFERENCE: P-LJ 5254
, CURRENT APPLICATION NUMBER: US/10/141,618
, CURRENT FILING DATE: 2002-05-07
, PRIOR APPLICATION NUMBER: US 60/289,233
, PRIOR FILING DATE: 2001-05-07
, PRIOR APPLICATION NUMBER: US 60/356,934
, PRIOR FILING DATE: 2002-02-12
, PRIOR APPLICATION NUMBER: US 09/388,221
, PRIOR FILING DATE: 1999-09-01
, NUMBER OF SEQ ID NOS: 15
, SOFTWARE: FastSeq For Windows Version 4.0
, SEQ ID NO 10

```

[illegible]

Qy 961 EAQVSCCLSPHLOVIAFGDENGAEILELVNRRIFQSRFOHKKTVWHIOFTADEKTLIS 1020
Db 907 EAQVSCCLSPHLOVIAFGDENGAEILELVNRRIFQSRFOHKKTVWHIOFTADEKTLIS 966
Qy 1021 SSDDAEIQVNNWQDKICIFLFGHGTVDKDFRLLKNSRLLSWSFDGTVKVMNIIITGNKEKD 1080
Db 967 SSDDAEIQVNNWQDKICIFLFGHGTVDKDFRLLKNSRLLSWSFDGTVKVMNIIITGNKEKD 1026
Qy 1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLPLPLHLRGNCGVCSAFSDVST 1140
Db 1027 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLPLPLHLRGNCGVCSAFSDVST 1086
Qy 1141 LLATGDDNGEIRIWNVSGELHLICAPLSEGAATHGCGWTDLCFSPDGKMLISAGGYIK 1200
Db 1087 LLATGDDNGEIRIWNVSGELHLICAPLSEGAATHGCGWTDLCFSPDGKMLISAGGYIK 1146
Qy 1201 WNVVTGSSQTFYNGNLKXIHVSPDFKTYVVDNLGILYILOTLE 1248
Db 1147 WNVVTGSSQTFYNGNLKXIHVSPDFKTYVVDNLGILYILOTLE 1194

RESULT 5

US-10-112-944-273
; Sequence 273, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Weng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Gao, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; TITLE OF INVENTION: Secreted Polypeptides
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt FL genes Version 5.0
; SEQ ID NO 273
; LENGTH: 1199
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-944-273

Query Match 90.7%; Score 6003.5; DB 15; Length 1199;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 1154; Conservative 0; Mismatches 0; Indels 139; Gaps 5;

Qy 1 MDKARNCLLOHREALEKDKITSYIMDHMSDGLFTITSEBEKVNEPTQOQRAAMLIKMI 60
|||||

Db 1 MDKARNCLLOHREALEKDKITSYIMDHMSDGLFTITSEBEKVNEPTQOQRAAMLIKMI 60
Qy 61 LKKNDSYVSPYNALLHEGYKDLAALLHDGIPVYSSSGKSDSVSGITSYVTVI.CEGVGP 120
Db 61 LKKNDSYVSPYNALLHEGYKDLAALLHDGIPVYSSSGS-----VTVI.CEGVGP 109
Qy 121 QRPVVFVTRKKLVNAIOOKLSKLGEGPWTIHHMAGCGKSVLAEEAVRDSHSLLEGCPGP 180
Db 110 QRPVVFVTRKKLVNAIOOKLSKLGEGPWTIHHMAGCGKSVLAEEAVRDSHSLLEGCPGP 169
Qy 181 GVHWVSGVKDKSGLLKNLQNLCTRLDQDSEFSQRLPLNIEEAKDRLRLMLRKHPRSLL 240
Db 170 GVHWVSGVKDKSGLLKNLQNLCTRLDQDSEFSQRLPLNIEEAKDRLRLMLRKHPRSLL 229
Qy 241 ILDDVMSWLVKAFDSQCQILLTTTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN 300
Db 230 ILDDVMSWLVKAFDSQCQILLTTTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN 289
Qy 301 MKKADLPEQAHSIIECKGSPVLVSLIGALLRDPNREWEYVYLKQIQKQFRIKSSSYD 360
Db 290 MKKADLPEQAHSIIECKGSPVLVSLIGALLRDPNREWEYVYLKQIQKQFRIKSSSYD 344
Qy 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKDVKVPKVLICILMDMETEEVEDILOEPVN 420
Db 345 YEALDEAMSISVEMLRREDIKDYITDLSILOKDVKVPK----- 382
Qy 421 KSLLFCDNRNGKSFYRYLHDLQVDFLTBKNSQLQDLHKIKITQFQRYHQPHTLSPDQSDC 480
Db 383 -----DLHKIKITQFQRYHQPHTLSPDQSDC 408
Qy 481 MYWYNFLAYHMASAKMKELCALMFSLDWIKAKTELVGPAHLIHEFVYRHLIHEKOCAN 540
Db 409 MYWYNFLAYHMASAKMKELCALMFSLDWIKAKTELVGPAHLIHEFVYRHLIHEK---- 464
Qy 541 SENFQEFLSLNGHLLGRPPENIVOLGLCEPSETSEVTVQQAQKQAEVDNGLYLEWINK 600
Db 465 -----PPENIVOLGLCEPSETSEVTVQQAQKQAEVDNGLYLEWINK 506
Qy 601 KNITNLSRLVVRPHTDAVYHACFSEDEGRIASCADKTLQ----- 640
Db 507 KNITNLSRLVVRPHTDAVYHACFSEDEGRIASCADKTLQQAASGSTEAGPLPGAA 566
Qy 641 -----VFAETGEKLEIKAHEDVLCAPSTDDRPFATC 675
Db 567 VAAQAAVADSGIPVLLGLDLEGGPPALVFAETGEKLEIKAHEDVLCAPSTDDRPFATC 626
Qy 676 SVDKKKVKNWSMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGSSDCFLKLDLNOKEC 735
Db 627 SVDKKKVKNWSMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGSSDCFLKLDLNOKEC 686
Qy 736 RNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLWDATSANERKSINVKQFFLNLEDPQE 795
Db 687 RNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLWDATSANERKSINVKQFFLNLEDPQE 746
Qy 796 DMEVIVKCCSWSDGARIMVAANKKIFLFDIHTSGLLGEIHTGHSTIYQCDPSPQNHLA 855
Db 747 DMEVIVKCCSWSDGARIMVAANKKIFLFDIHTSGLLGEIHTGHSTIYQCDPSPQNHLA 806
Qy 856 VVALSOYCVBELWNTDSKVDACRGLHSWVHGFSPGSSFLTSSDDOTIRLWETKKVC 915
Db 807 VVALSOYCVBELWNTDSKVDACRGLHSWVHGFSPGSSFLTSSDDOTIRLWETKKVC 866
Qy 916 KNSAVMLKQEVVDVVFQENEVNVLAVDHRRRLQLINGRTQIDYILTEAQVSCCLSPHLOQ 975
Db 867 KNSAVMLKQEVVDVVFQENEVNVLAVDHRRRLQLINGRTQIDYILTEAQVSCCLSPHLOQ 926
Qy 976 IAFDENGAEIIELELVNRRIFQSRFOHKKTVWHIOFTADEKTLISSDDAEIQVNNWOLD 1035
Db 927 IAFDENGAEIIELELVNRRIFQSRFOHKKTVWHIOFTADEKTLISSDDAEIQVNNWOLD 986
Qy 1036 KCIFLRGHOETVKDFRLLKNSRLLSWSFDGTVKVMNIIITGNKEKDFVCHQGTVLSCDISH 1095
Db 987 KCIFLRGHOETVKDFRLLKNSRLLSWSFDGTVKVMNIIITGNKEKDFVCHQGTVLSCDISH 1046

QY 1096 DATKSTSDAKTAKIMSDLLPLHELHNGCVRCFAFSDSTLLATGDNGEIRIWN 1155
 DB 1047 DATKSTSDAKTAKIMSDLLPLHELHNGCVRCFAFSDSTLLATGDNGEIRIWN 1106
 QY 1156 VSNGLLHLHLCAPLSEGAATHGHWTDLCFSPDGKMLISAGYIKWNNVVTGSSOTFTY 1215
 DB 1107 VSNGLLHLHLCAPLSEGAATHGHWTDLCFSPDGKMLISAGYIKWNNVVTGSSOTFTY 1166
 QY 1216 NGTNLKKHVSPPDKTYVVDNGLIYLLOLLE 1248
 DB 1167 NGTNLKKHVSPPDKTYVVDNGLIYLLOLLE 1199

RESULT 6

US-10-112-944-272
 ; Sequence 272, Application US/10112944
 ; Publication No. US20040048249A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Gezhi
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Ghosh, Malabika
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wang, Zhiwei
 ; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
 ; TITLE OF INVENTION: Secreted Polypeptides
 ; FILE REFERENCE: 805A
 ; CURRENT APPLICATION NUMBER: US/10/112, 944
 ; PRIOR FILING DATE: 2002-03-28
 ; PRIOR FILING DATE: 2000-01-21
 ; PRIOR FILING DATE: 2000-01-21
 ; PRIOR FILING DATE: 2000-01-25
 ; PRIOR FILING DATE: 2000-02-03
 ; PRIOR FILING DATE: 2000-02-03
 ; PRIOR FILING DATE: 2000-02-28
 ; PRIOR FILING DATE: 2000-03-07
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR FILING DATE: 2000-04-18
 ; PRIOR FILING DATE: 2000-05-18
 ; NUMBER OF SEQ ID NOS: 924
 ; SOFTWARE: pt_FL_genes Version 5.0
 ; SEQ ID NO 272
 ; LENGTH: 1526
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-112-944-272

Query Match 89.6%; Score 5933.5; DB 15; Length 1526;
 Best Local Similarity 89.1%; Pred. No. 0;
 Matches 1141; Conservative 0; Mismatches 0; Indels 139; Gaps 5;
 QY 14 EALEKDITSYMDHMSDGLTISBEEKVRNEPTQOQRAAMLKMLKKDNDSYVSFYN 73
 DB 341 EALEKDITSYMDHMSDGLTISBEEKVRNEPTQOQRAAMLKMLKKDNDSYVSFYN 400
 QY 74 ALHLEGYKDLAALLHGDGIPVSSSSGKDSVSGITSYRTVLCGGVFPQRPVFTTRKLV 133
 DB 401 ALHLEGYKDLAALLHGDGIPVSSSS-----VTVLCGGVFPQRPVFTTRKLV 449
 QY 134 NAIQOKLSKLGEPGWTTIHGAGCGKSVLAEEAVRDSHLLLEGCFPGGVHWSVGKQDKS 193

DB 450 NAIQOKLSKLGEPGWTTIHGAGCGKSVLAEEAVRDSHLLLEGCFPGGVHWSVGKQDKS 509
 QY 194 GLMKLQNLCTRLDQDESFQRLPLNIEEAKDRILMLRKHPRSLILLDDVWDSWVLA 253
 DB 510 GLMKLQNLCTRLDQDESFQRLPLNIEEAKDRILMLRKHPRSLILLDDVWDSWVLA 569
 QY 254 FDSQCOILLTRTKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVNMKADLPQAHSI 313
 DB 570 FDSQCOILLTRTKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVNMKADLPQAHSI 629
 QY 314 IKECKGSPVLSIGALLRDFNRWEYILKOLQKQFRIRKSSSYDYEALDEAMSISVE 373
 DB 630 IKECK-----VSLIGALLRDFNRWEYILKOLQKQFRIRKSSSYDYEALDEAMSISVE 684
 QY 374 MLREDIKDYITDLSILQKQVPTKVLCLMDMETEVEDILOEFVNKSLPCDRNGKSF 433
 DB 685 MLREDIKDYITDLSILQKQVPTK----- 709
 QY 434 RYLLHDLQVDFLTERKNCSQLDLHKKIITQFYHQPHTLSPQEDCMYWNFLAYHMAS 493
 DB 710 -----DLHKKIITQFYHQPHTLSPQEDCMYWNFLAYHMAS 748
 QY 494 AKMHKELCALMFSLDWIKARTELVGPAHLIHEFVEYRHLDEKDCAVSENFOFFLSLNGH 553
 DB 749 AKMHKELCALMFSLDWIKARTELVGPAHLIHEFVEYRHLDEK----- 791
 QY 554 LLGRQPPFNIVQLGCEPETSEVYQOAKLOAKQVVDNGLYLEWINKNITNLSRLVVRP 613
 DB 792 -----PPFNIVQLGCEPETSEVYQOAKLOAKQVVDNGLYLEWINKNITNLSRLVVRP 846
 QY 614 HTDAVYHACFSEDOGRIASGADKTLQ----- 640
 DB 847 HTDAVYHACFSEDOGRIASGADKTLQLOARASGSTRQAGPLPPGAVAQAADSGIP 906
 QY 641 -----VFKAETGEKLEIKAHDEVILCASTDDRIATCSVDKVKIWNST 688
 DB 907 VLGDLEGPPALVFKAETGEKLEIKAHDEVILCASTDDRIATCSVDKVKIWNST 966
 QY 689 GELVHTYDEHSEQVNCCHFTNSHHLLLATGSSDCFLKLDLNQKCRNTPMGHTNSVNH 748
 DB 967 GELVHTYDEHSEQVNCCHFTNSHHLLLATGSSDCFLKLDLNQKCRNTPMGHTNSVNH 1026
 QY 749 CRFSPDDKLLASCSADGTLKLDWATSANERKSINVKQFFLNLEDPOEDMEVIVKCCWSA 808
 DB 1027 CRFSPDDKLLASCSADGTLKLDWATSANERKSINVKQFFLNLEDPOEDMEVIVKCCWSA 1086
 QY 809 DGARIMVAANKIIFLFDIHTSGLLGEIHTGHSTIQYCDFSPQNLHVALSVQYCVELWN 868
 DB 1087 DGARIMVAANKIIFLFDIHTSGLLGEIHTGHSTIQYCDFSPQNLHVALSVQYCVELWN 1146
 QY 869 TDSRSKVADCRGHLVWVGWVFPDSSFLTSDDQIRLWETKVKCKNSAVMLKQEVVDV 928
 DB 1147 TDSRSKVADCRGHLVWVGWVFPDSSFLTSDDQIRLWETKVKCKNSAVMLKQEVVDV 1206
 QY 929 VFOENVMVLAVDHIRRLQILINGRTQIDYLTAEQVSCCLSPHLOYIAFGDENGAIEL 988
 DB 1207 VFOENVMVLAVDHIRRLQILINGRTQIDYLTAEQVSCCLSPHLOYIAFGDENGAIEL 1266
 QY 989 ELVNNRIFQSRFOHKKVWHIOFTADEKTLISSDDAEIOVNNWQDKCIFLGHQETVK 1048
 DB 1267 ELVNNRIFQSRFOHKKVWHIOFTADEKTLISSDDAEIOVNNWQDKCIFLGHQETVK 1326
 QY 1049 DFRLLKNSRLLSFPGTVKWNITGKNEKDFVCHOGTVLSCDISHDATKFSSTADKT 1108
 DB 1327 DFRLLKNSRLLSFPGTVKWNITGKNEKDFVCHOGTVLSCDISHDATKFSSTADKT 1386
 QY 1109 AKIWSFDLLPLHELHNGCVRCFAFSDSTLLATGDNGEIRIWNVNGELLHLCAPL 1169
 DB 1387 AKIWSFDLLPLHELHNGCVRCFAFSDSTLLATGDNGEIRIWNVNGELLHLCAPL 1446
 QY 1169 SEEGAATHGCVTDLCFSPDGKMLISAGYIKWNNVVTGSSOTFTYNTNKLKIHVSPD 1228

Db 1447 SEEGNATHGGWTDLCFSPDGMKLISAGYIKWNNVVTGESSQFTYNGTNLKKIHVSPD 1506
Qy 1229 FKTYVTVDNLGLYILQTL 1248
Db 1507 FKTYVTVDNLGLYILQTL 1526

RESULT 7

US-10-014-269-31
; Sequence 31, Application US/10014269
; Publication No. US20020127673A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Ogura, Yasunori
; TITLE OF INVENTION: NOD2 Nucleic Acids and Proteins
; FILE REFERENCE: UM-06645
; CURRENT APPLICATION NUMBER: US/10/014,269
; PRIORITY FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 31
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-014-269-31

Query Match 24.0%; Score 1589; DB 13; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.8e-118;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 123 PVVFTVRKKLVNAIQOKLSKLKGPWGVTIHGMAGCGKSVLAABAVRDHSLLEGCPGGV 182
Db 1 PVVFTVRKKLVNAIQOKLSKLKGPWGVTIHGMAGCGKSVLAABAVRDHSLLEGCPGGV 60
Qy 183 HWSVGKQDKSGLMKLQNLCTRLDQDSFSQRLPLNIEEAKDRILMLRKHPRSLIL 242
Db 61 HWSVGKQDKSGLMKLQNLCTRLDQDSFSQRLPLNIEEAKDRILMLRKHPRSLIL 120
Qy 243 DDVWDSWVLKAFDSCQILLTTRDKSVTDSVMGPKYVVPVSSIGKEKGLSILFVNMM 302
Db 121 DDVWDSWVLKAFDSCQILLTTRDKSVTDSVMGPKYVVPVSSIGKEKGLSILFVNMM 180
Qy 303 KADLPEQAHSIIKECKGSPVLSLIGALLRDPFNNRWEYLLKQLNQKQPKRIRKSSSYDYE 362
Db 181 KADLPEQAHSIIKECKGSPVLSLIGALLRDPFNNRWEYLLKQLNQKQPKRIRKSSSYDYE 240
Qy 363 ALDEAMSI SVMELREDIKDYTTDLSILQKQVKTVCILWDMETEVEEDILQEFVNKS 422
Db 241 ALDEAMSI SVMELREDIKDYTTDLSILQKQVKTVCILWDMETEVEEDILQEFVNKS 300
Qy 423 LLFCDRNG 430
Db 301 LLFCDRNG 308

RESULT 8

US-10-002-974-31
; Sequence 31, Application US/10002974
; Publication No. US20020197616A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Ogura, Yasunori
; APPLICANT: Cho, Judy
; APPLICANT: Nicolae, Dan L
; APPLICANT: Bonen, Denise
; TITLE OF INVENTION: NOD2 Nucleic Acids and Proteins
; FILE REFERENCE: UM-06646
; CURRENT APPLICATION NUMBER: US/10/002,974
; PRIORITY FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: Patent in version 3.1

; SEQ ID NO 31
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-974-31

Query Match 24.0%; Score 1589; DB 13; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.8e-118;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 123 PVVFTVRKKLVNAIQOKLSKLKGPWGVTIHGMAGCGKSVLAABAVRDHSLLEGCPGGV 182
Db 1 PVVFTVRKKLVNAIQOKLSKLKGPWGVTIHGMAGCGKSVLAABAVRDHSLLEGCPGGV 60
Qy 183 HWSVGKQDKSGLMKLQNLCTRLDQDSFSQRLPLNIEEAKDRILMLRKHPRSLIL 242
Db 61 HWSVGKQDKSGLMKLQNLCTRLDQDSFSQRLPLNIEEAKDRILMLRKHPRSLIL 120
Qy 243 DDVWDSWVLKAFDSCQILLTTRDKSVTDSVMGPKYVVPVSSIGKEKGLSILFVNMM 302
Db 121 DDVWDSWVLKAFDSCQILLTTRDKSVTDSVMGPKYVVPVSSIGKEKGLSILFVNMM 180
Qy 303 KADLPEQAHSIIKECKGSPVLSLIGALLRDPFNNRWEYLLKQLNQKQPKRIRKSSSYDYE 362
Db 181 KADLPEQAHSIIKECKGSPVLSLIGALLRDPFNNRWEYLLKQLNQKQPKRIRKSSSYDYE 240
Qy 363 ALDEAMSI SVMELREDIKDYTTDLSILQKQVKTVCILWDMETEVEEDILQEFVNKS 422
Db 241 ALDEAMSI SVMELREDIKDYTTDLSILQKQVKTVCILWDMETEVEEDILQEFVNKS 300
Qy 423 LLFCDRNG 430
Db 301 LLFCDRNG 308

RESULT 9

US-10-314-506-31
; Sequence 31, Application US/10314506
; Publication No. US20030175762A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Ogura, Yasunori
; TITLE OF INVENTION: Modulators of NOD2 Signaling
; FILE REFERENCE: UM-06984
; CURRENT APPLICATION NUMBER: US/10/314,506
; PRIORITY FILING DATE: 2002-12-09
; PRIORITY FILING DATE: 2001-10-26
; PRIORITY FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 31
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-314-506-31

Query Match 24.0%; Score 1589; DB 14; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.8e-118;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 123 PVVFTVRKKLVNAIQOKLSKLKGPWGVTIHGMAGCGKSVLAABAVRDHSLLEGCPGGV 182
Db 1 PVVFTVRKKLVNAIQOKLSKLKGPWGVTIHGMAGCGKSVLAABAVRDHSLLEGCPGGV 60
Qy 183 HWSVGKQDKSGLMKLQNLCTRLDQDSFSQRLPLNIEEAKDRILMLRKHPRSLIL 242
Db 61 HWSVGKQDKSGLMKLQNLCTRLDQDSFSQRLPLNIEEAKDRILMLRKHPRSLIL 120
Qy 243 DDVWDSWVLKAFDSCQILLTTRDKSVTDSVMGPKYVVPVSSIGKEKGLSILFVNMM 302
Db 301 LLFCDRNG 308

Db 121 DVWDSWVLKAFDSQCLLTTRDKSVTDSVMGPKVYVPVSSLGKEKGLLEILSLFVNNK 180
 Qy 303 KADLPQAHSIKECKGSPVWSLIGALLRDPFNRWEYLLKOLQNKQPKIRKSSSYDYE 362
 Db 181 KADLPQAHSIKECKGSPVWSLIGALLRDPFNRWEYLLKOLQNKQPKIRKSSSYDYE 240
 Qy 363 ALDEAMSIIVMLREDIKDYDYLILQDKVKVPTKVLGCIILWDMETBEVEDILQEFVNS 422
 Db 241 ALDEAMSIIVMLREDIKDYDYLILQDKVKVPTKVLGCIILWDMETBEVEDILQEFVNS 300
 Qy 423 LLFCDRNG 430
 Db 301 LLFCDRNG 308

RESULT 10
 US-10-001-254-29
 ; Sequence 29, Application US/10001254
 ; Publication No. US20030049702A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, John C.
 ; APPLICANT: Godzik, Adam
 ; APPLICANT: Pawlowski, Krzysztof
 ; APPLICANT: Fiorentino, Loredana
 ; APPLICANT: Lee, Sug Hyung
 ; APPLICANT: Roth, Wilfred
 ; APPLICANT: Stenner-Liewen, Frank
 ; TITLE OF INVENTION: No. US20030049702A1el Death Domain Proteins
 ; FILE REFERENCE: P-LJ 5037
 ; CURRENT APPLICATION NUMBER: US/10/001,254
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: 60/301,889
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: 09/715,893
 ; PRIOR FILING DATE: 2000-11-17
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 29
 ; LENGTH: 142
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; US-10-001-254-29

Query Match 11.3%; Score 748; DB 14; Length 142;
 Best Local Similarity 100.0%; Pred. No. 1.4e-51;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 142 KLKGEFGWTHIGWAGCKSVLAEEAVRDHSILLEGCFGGVHWVSVGKQKSGLLMKLQN 201
 Db 1 KLKGEFGWTHIGWAGCKSVLAEEAVRDHSILLEGCFGGVHWVSVGKQKSGLLMKLQN 60
 Qy 202 LCTRLDQDESORPLNIEAKDRILMLKHPRSLLILDDVWDSWVLKAFDSQCLLT 261
 Db 61 LCTRLDQDESORPLNIEAKDRILMLKHPRSLLILDDVWDSWVLKAFDSQCLLT 120
 Qy 262 LTRDKSVTDSVMGPKVYVPVE 283
 Db 121 LTRDKSVTDSVMGPKVYVPVE 142

RESULT 11
 US-10-369-493-12913
 ; Sequence 12913, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 12913
 ; LENGTH: 529
 ; TYPE: PRT
 ; ORGANISM: Aspergillus nidulans
 ; US-10-369-493-12913
 Query Match 10.7%; Score 711; DB 14; Length 529;
 Best Local Similarity 29.0%; Pred. No. 1e-47;
 Matches 183; Conservative 93; Mismatches 242; Indels 112; Gaps 11;
 Qy 614 HTDAVTHACFSEDSQRIASCGADKTLQVFKAEIKLEIKAHEDVLCFCARSTDRFTA 673
 Db 4 HLGVPESVVFSPDGKQLVSGSYDDTVKIWDPATGELLQNDGHSVLESFSPDGKLLA 63
 Qy 674 TCSVDKVKIWNMTGELVHTYDEHSEOVNCHFTNSHHLLATGSSDCLFLKLDLQNK 733
 Db 64 SGSYDNTIDLWDSPTGELLQTFEGHPHSIWSVAFAPDKE--LASASDDSTIKINDLATG 121
 Qy 734 ECRNFMFGHTNSVNHCRFSPDDKLLASCSADGTLKLDATSANERKSINVKQFFLNEDP 793
 Db 122 ELQQLDHSQSRSVAFSPDGKLLASSLSDSTIKVWNPATGELQQLS----- 169
 Qy 794 QDEMEVIVKCCSWASADGARIMVAA--KNKIFLFDIHTSGLLGEIHTGHSTIOYCDSPON 852
 Db 170 --EGRSGWVKSVAFPDGKLLASGSEKNTVKLWNPATGELLQTL--GHSSOSVRSVAFSPDG 227
 Qy 853 HLAVALSOYCVELMNTDSRSKVADCRHLSVHGVWVFPDGSFILTSSDDOTIRLWETK 912
 Db 228 KQLASSSDSTIKLWNSITGELQQTQFKGHDLWIRAVAFSPDGKHLVSGSDDNITIKLW--- 284
 Qy 913 KVCNSAVMLKQEVVDVFOENEVMVLAVDHIRRLQINGRTGQIDYLTEAQVSCCCLSPH 972
 Db 285 ----- 284
 Qy 973 LQYIAGDENGAEILELVNRRIFQSRPHKKTVMHIQFTADEKTLISSDDAEIQVWN- 1031
 Db 285 -----DLATSELOQSLSDHSRSVHAFVAFSPDDKQLASSSLSDSTIKLWDS 328
 Qy 1032 --WOLDKCFILRGHQTVDKFRLLKNSRL--SWSFDTGTVKVMNIITGNKEKDFVCHQTV 1088
 Db 329 ATGELQRT--LEGHSQGVRSVTFSPDGKLLASNSYDGTIKLWNPATGELLQTLTGRSDWV 386
 Qy 1089 LSCDISHDATKFSSTADKTAKIWSFDLILLPLHELRGHNGVRCSCAFSVSDSTLLATGDN 1148
 Db 387 DSVAFSPDGKQLASGYDSTIKLWDSATGELLQTLLEGHSDRIQSVVFPDGLKASGPTN 446
 Qy 1149 GEIRVNVSGELLHLICAPLSEGAHGHGVTDLCSFDGKMLISA--GGYIKWNVVT 1206
 Db 447 KPAKFWDFATGELQIF-----EG--HKKWVESVAFSPDGKLLASSSYGETIKLWDPT 498
 Qy 1207 GSSQQTFTYTGNTLKKIHVSDFKTYVTVD 1236
 Db 499 GELLQTLNDPDESAGSVAFSPDGNRLASVD 528

RESULT 12
 US-10-369-493-20185
 ; Sequence 20185, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B

```

; CURRENT APPLICATION NUMBER: US/10/369,493
;
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
;
; SEQ ID NO 20185
; LENGTH: 610
; TYPE: PRT
; ORGANISM: No. US20030233675A1toc punctif.
US-10-369-493-20185

```

Query Match	10.4%;	Score 690.5;	DB 14;	Length 610;
Best Local Similarity	28.7%;	Pred. No. 5.7e-46;		
Matches 187;	Conservative 106;	Mismatches 273;	Indels 85;	Gaps 14;
Qy	614	HTDAVHACFSEDCQRIASCGADKTLOVFKAEKTGKELLEIKAHDEVLCAAFSTDDRFIA	673	
Db	5	HAKEVQGISFSPDGKMLASDDNTVKLWDTTTGKEIKTLTGHTNSVLGTSFSPDGKMLA	64	
Qy	674	TCSVDKKVKNWSMTGELVHTYDBHSBOVNCCHFTNSHHLLLATGSSDCFLKLWDLNNQK	733	
Db	65	SASSDNTVKLWDTTTGKEIKTLTGHTNSVLGISF--SPDGKMLASADNTVKLWDTTTG	122	
Qy	734	ECRNTMEGHTNSVNHCRFSPDDKLLASCSDGTLKLWDATSANERKSI NVKQFLNLEDP	793	
Db	123	KEIKTLTGHRNSVFGISFSPDGKMLASADNTVKLWDTTTGKEIKTL-----	170	
Qy	794	QEDMEVIVKCCSWSADGARIVAAAKNIFLFDIHTSGLLGGEIHTGHSHTIQCYDFSPONH	853	
Db	171	-----TGHRNSVFGISFSPDGK 187		
Qy	854	LAVVALSOYCVELWNTDSRSKVADCRGHLSMVHGWMESPSSFLTSSDDOITRLMWTK-	912	
Db	188	MLASASPDNTVKLWDTTTGKEIKTLTGHRNSVNDISFSPDGKMLASADNTVKLWDTTT	247	
Qy	913	-KVCKNSAVMLKQEVVDVVOFQENEVMLAVDHIRELQLINGRTG-QIDYLT--EAQVSCCC	968	
Db	248	GKEIKTLTGHRNSVNDISFSPNGKMLASASPDNTVKLWDTTTGKEIKTLTGHTNSVNDIS	307	
Qy	969	LSPHLQVIAFGDENGATEIELLVNRRIFQSRFQHKTVVHIQFTADSEKTLISSDDAEIQ	1028	
Db	308	FSPDGKMLASAGDNTVKLWDTTTGKEIKTLTGHRNSVNDISFSPDGKMLASAGDNTVK	367	
Qy	1029	VANWQLDKCT-FLRGHGETVKDPRLLKNRLL--SWSFDGTVKVWNITGNKEKDFVCHQG	1086	
Db	368	LWDTTTGKEIKTLTGHTNSVNGISFSPDGKMLASAGDKTVKLVWDTTTGKEIKTLTGHTN	427	
Qy	1087	TVLSCDISHDATKPSSTSAKTAKIWSFDLLLPLHELRGHNGCVKCSAFSVDSLTLAGD	1146	
Db	428	SVANGISFSPDGKMLASAGDKTVKLVWDTTTGKEIKTLTGHTNSVNGISFSPDGKMLAS	487	
Qy	1147	DNGEIRLWN-VSNGELLHLCAPLSEEGAAHTGGWVTDLCFSPDGKMLISAG--GYIKWN	1203	
Db	488	SDNTVKLWDTTTGKTL-----KTLTGHTNSVNGISFSPDGKMLASADNTVKLWD	539	
Qy	1204	VWTGESSQTFYTNGTN-LKKIHVSPDFK--TYVTVDN-----LGILYILQ	1245	
Db	540	TTTGKEIKTL-TGHTNVMYGISFSPDGKMLASADNTVKLWRLDYFLLQ	509	

RESULT 13

```

RES001 13
US-10-369-493-20242
; Sequence 20242, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B

```


Db 46 ASGSGKSS-----VQAGLIPQ-----LRQKHIPNSQWGIKTIIR--PGVNPLEA 89
 Qy 155 MAGCKSVLAABAVRDHSLLEGCF-----PGGVHVSQKQDQXGGLM--KLQNLCTRL 206
 Db 90 LAR-----KLGEWGETHLLIEGMLHQGVFVYMLRSLPQGVTVLVDQFEELEFLAPT 144
 Qy 207 DQDESQRLPLNIBEAQRDLRLMLRKHPRLSLLILDDVWDSWVLKAFDSQCQILLTRD 266
 Db 145 DR-EUFLLELLGAVQAGDRFKLI-----TLRA----- 172
 Qy 267 KSVTDSVMGPKYVVPVVESSLGKEGLEILSLFVNMMKADLPQAHSLIKECK--GSPLVW 324
 Db 173 -----DFAFCLEVPALAEALQVAVLV-PPKLSLDYRRVILHPAQOQVGLQVEA 221
 Qy 325 SLIGALLR-----DFPNRWEYLLKQONKQKFKIRKSSSYD----- 360
 Db 222 ELVEVLLRELNGSVGDLP-LLEFVLEQLWQORTAGKLTLSYQEQELGGIKGALERSCOGV 280
 Qy 361 YEALDEAMS-----ISVEMLREDIKO-----YVTDL-----SILQKOVKVPKVL 401
 Db 281 YESLPPQOQCAKWIFLSLTLQGEGETEDTRRIYKSDLIYKYPAGLVEQTLNVLNNAKL 340
 Qy 402 ILWDMETEVEDILQEFVNKSLFCDRNGKSPRYLHDLQVDFLTEKNCSQLODLHKKII 461
 Db 341 VVINLEAE-----IEAQKSPSPPTPNLSTPFVTVVEVAHEILIRHWSL 384
 Qy 462 TOFORYHQPHTLSPDQ--EDCMYVY-----NFLAYHMASAKMKELCALMFSLDWIRAK 513
 Db 385 RWMLEENRDLRKQRIHACQLWQOSGQADFL--LQGARLAEAEIDIIYVWTDLGAD 441
 Qy 514 T-ELVGP--AHLIHEFVEYRHILDEKCAV-----SENQOEF 548
 Db 442 VQEFIGACLAERKHQOQAKRNLKQARAVALSVLGIASVFGGLAYWQGREAQFREIA 501
 Qy 549 SLN-----GHLGRQPPNIVOLGCEPETSEVYQAKLQAKQOEVNDGMLYLEWINKNIT 604
 Db 502 ALNSSQANLLSHQQLAALI-----ASLKAQOVNHHVIAVNNLKLATV 546
 Qy 605 NLSRLV-----VRPHTDAVYHACFSEDGQRIASCGADKTLOVFAETGEKLEIKAH 656
 Db 547 TLQOALPEMOBRNLEGHKGDISISIRDGQTIASGSLDKTIKLMSRD-GRLPFTLNGH 605
 Qy 657 EDEVLCFAFSTDDRIATCSVDKVKIWNMTGELVHYTDEHSEOVNCHFTNSSHLLL 716
 Db 606 EDVYVSFSFDPGQTIASGSDTKIKLWQTSDTGLTKITITGHEQTVNNVYFSPDGKN--L 663
 Qy 717 ATGSSDCFLKLWDLNQKCRNTMGHTNSVNHCFSPDDKLLASCADGTLKJWDATSAN 776
 Db 664 ASASDHSIKLWDTTSGQLLMTLTGHSAGVITVRFSPDQTTAAGSEDKTVKLWHRQDGK 723
 Qy 777 ERKSINVKQFFLNLEBDQEDMEVIVKCCSWSADGARIMVAANKIIFLFDIHTSGLLGEIH 836
 Db 724 LLKTLNGHQDWV-----SLSFSPDGKTLASASADKTIKLWRIADGKLVKTL 770
 Qy 837 TGHSTIOYCFPSQNHIAVALSOYCVELWNTDSRSKVADCRHLSWVHGMFSPDGSS 896
 Db 771 KGHDSVWVDFNSPDGKAIASASRDNTIKLWNRHG-IELETFTGHSGGVYAVNFLPDSNI 829
 Qy 897 FLTSSDDOTIRLWE-----TKVKCNKSAVMLKQEVDFVQENEVWVLAVDHIRRLQIN 950
 Db 830 IASASLONTIRLWQRLPLSPLEVLNAGNSGVA-----VSFLHDGSIITAGADGNIQLWH 884
 Qy 951 GRTQIDYLTBAQVSCCLLS--PHLQYIAFGDENGAIIEILVNNRIFQSRFOHKKTVWH 1008
 Db 885 SQDGSLLKTLPCNKAIYGISFTPPQGDLIASANADKTVKIWRVROGKALKTLIGHDNEVVK 944
 Qy 1009 IOFTADEKTLISSDDABIQVNNWQDKCIF---LRGHOETVKDPRLLKNSRL--SNSFD 1064
 Db 945 VNFSPDGKTLASASRDNTVKLWNVSDGK--FKTKLKGHTDEVFWVSFSPDGKIIASAD 1002
 Qy 1065 GTVKVWNIITGNKEKDFVCHQGTVLSCDISHDATKFSSTADKTIKW-SFDLLPLHEL 1123
 Db 1003 KTIRLWDSFGNLIKSLPAHNDLVYSVNFNPDGMSLASTSADKTVKLWRSHDGL-LHTF 1061

Qy 1124 RHNGCVRCSAFSDVSTLLATGDNGEIRIWNVSGELLHLHCAPLSEGAATHGGWVTDL 1183
 Db 1062 SGHSNVVYSSSPSPDGRYIASASEDKTVKIWDIG---HLLTTLTPQ-----HQAGVMSA 1112
 Qy 1184 CFSPDGKMLIS 1194
 Db 1113 IFSPDGKTLIS 1123

Search completed: January 27, 2005, 18:53:21
 Job time : 178 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 18:32:45 ; Search time 27 Seconds
(without alignments)
3065.365 Million cell updates/sec

Title: US-10-646-396-2

Perfect score: 6619

Sequence: 1 MDKARNCLLQHRLEAKDI.....FKTVVVDNLGILYLQTL 1248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6352.5	96.0	1205	3	US-09-092-508-16
2	6352.5	96.0	1205	3	US-09-435-115-16
3	6282	94.9	1194	3	US-09-092-508-2
4	6282	94.9	1194	3	US-09-435-115-2
5	6282	94.9	1194	3	US-09-069-023-26
6	6282	94.9	1194	4	US-09-098-310-2
7	6282	94.9	1194	4	US-09-538-092-825
8	408	6.2	2627	2	US-08-751-189-3
9	408	6.2	2627	2	US-09-060-836-3
10	408	6.2	2627	3	US-09-184-445-3
11	399.5	6.0	2629	2	US-08-751-189-4
12	399.5	6.0	2629	2	US-09-060-836-4
13	399.5	6.0	2629	3	US-09-184-445-4
14	322	4.9	251	4	US-09-291-170A-13
15	322	4.9	251	4	US-09-724-884-13
16	322	4.9	251	4	US-09-724-592-13
17	316	4.8	514	1	US-08-190-802A-66
18	316	4.8	514	3	US-08-477-346-66
19	316	4.8	514	3	US-08-473-089-66
20	316	4.8	514	4	US-08-487-072A-66
21	312	4.7	409	1	US-08-190-802A-51
22	312	4.7	409	3	US-08-477-346-51
23	312	4.7	409	3	US-08-473-089-51
24	312	4.7	409	4	US-08-487-072A-51
25	311.5	4.7	409	2	US-08-283-917-3
26	311.5	4.7	409	2	US-08-961-716-3
27	311.5	4.7	409	4	US-09-538-092-1119

28	311.5	4.7	410	2	US-08-283-917-9	Sequence 9, Appli
29	311.5	4.7	410	2	US-08-961-716-9	Sequence 9, Appli
30	309	4.7	318	1	US-08-190-802A-33	Sequence 33, Appli
31	309	4.7	318	3	US-08-477-346-33	Sequence 33, Appli
32	309	4.7	318	3	US-08-473-089-33	Sequence 33, Appli
33	309	4.7	318	4	US-08-487-072A-33	Sequence 33, Appli
34	307.5	4.6	316	4	US-09-828-310-12	Sequence 12, Appli
35	297.5	4.5	640	4	US-09-177-165A-30	Sequence 30, Appli
36	294	4.4	540	4	US-09-213-888-7	Sequence 7, Appli
37	294	4.4	540	4	US-09-213-888-10	Sequence 10, Appli
38	294	4.4	540	4	US-09-328-877D-7	Sequence 7, Appli
39	294	4.4	540	4	US-09-328-877D-10	Sequence 10, Appli
40	294	4.4	545	4	US-09-213-888-6	Sequence 6, Appli
41	294	4.4	545	4	US-09-328-877D-6	Sequence 6, Appli
42	294	4.4	553	4	US-09-213-888-5	Sequence 5, Appli
43	294	4.4	553	4	US-09-328-877D-5	Sequence 5, Appli
44	294	4.4	559	4	US-09-213-888-9	Sequence 9, Appli
45	294	4.4	559	4	US-09-328-877D-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-092-508-16
; Sequence 16, Application US/09092508
; Patent No. 6291643
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6291643west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/092,508
; FILING DATE: 05-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,807
; FILING DATE: 05-JUN-1997
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-09-092-508-16

Query Match 96.0%; Score 6352.5; DB 3; Length 1205;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1205; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

Qy 1 MDKARNCLLOHREALEKDIKTSYIMDMHISDGLFTTISEEEKVNEPTQOORAAMLIKMI 60
 Db 1 MDKARNCLLOHREALEKDIKTSYIMDMHISDGLFTTISEEEKVNEPTQOORAAMLIKMI 60
 Qy 61 LKKNDSVSVFYNALLHEGYKDLAALLHDGIIPWSSSSGKDSVSGITSYVVRTVLCGGVP 120
 Db 61 LKKNDSVSVFYNALLHEGYKDLAALLHDGIIPWSSSSGKDSVSGITSYVVRTVLCGGVP 120
 Qy 121 QRPVVVTRKLVNAIOOKLSKLGKPGWVTHGMAGCKSVLAABAVRSHLSLECCFPG 180
 Db 121 QRPVVVTRKLVNAIOOKLSKLGKPGWVTHGMAGCKSVLAABAVRSHLSLECCFPG 180
 Qy 181 GYHWSVQKQSGILMKLONLCTRLDDESFORSPLNIEEAKORILMLKXPHRSLL 240
 Db 181 GYHWSVQKQSGILMKLONLCTRLDDESFORSPLNIEEAKORILMLKXPHRSLL 240
 Qy 241 ILDDVMSVWLKAFDSQOILLTRDKSVTDVSMGPKVVPVSSSLGKGLBILSLFVN 300
 Db 241 ILDDVMSVWLKAFDSQOILLTRDKSVTDVSMGPKVVPVSSSLGKGLBILSLFVN 300
 Qy 301 MKKADLPQAHSSIIKECKGSPVLSLIGALLRDPFNRWEYLYKQONKQKFRIRKSSYD 360
 Db 301 MKKADLPQAHSSIIKECKGSPVLSLIGALLRDPFNRWEYLYKQONKQKFRIRKSSYD 360
 Qy 361 YEALDEAMSSIVEMLRDIDYVTDLSILQKQVKTVCILWDMETEEVEDILOEFVN 420
 Db 361 YEALDEAMSSIVEMLRDIDYVTDLSILQKQVKTVCILWDMETEEVEDILOEFVN 420
 Qy 421 KSLLCFDRNGKSFYRLHDLQVDFLTKNCSQDLHKKIITQFORVHQPHTLSPQEDC 480
 Db 421 KSLLCFDRNGKSFYRLHDLQVDFLTKNCSQDLHKKIITQFORVHQPHTLSPQEDC 480
 Qy 481 MYWYFLAYHMASAKMKECALMFSIDWKATKELVGPALHIEFEVYRHLDEXCAV 540
 Db 481 MYWYFLAYHMASAKMKECALMFSIDWKATKELVGPALHIEFEVYRHLDEXCAV 540
 Qy 541 SENFOEFLSLNGHILGROPENIVOLGCEPTESEVYQOAKLOKQOVDMGMLYLEWINK 600
 Db 541 SENFOEFLSLNGHILGROPENIVOLGCEPTESEVYQOAKLOKQOVDMGMLYLEWINK 600
 Qy 601 KMITLSRLVPRHTDAVTHACSEDOQRIASCADKTLQVFKARTGEKLEIKAHEDV 660
 Db 601 KMITLSRLVPRHTDAVTHACSEDOQRIASCADKTLQVFKARTGEKLEIKAHEDV 660
 Qy 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLLATGS 720
 Db 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLLATGS 720
 Qy 721 SDCFLKLDLQKQECRNTMFGHTNSVNHCRFSPDDKLLASCSDGTLKLDATSANERKS 780
 Db 721 SDCFLKLDLQKQECRNTMFGHTNSVNHCRFSPDDKLLASCSDGTLKLDATSANERKS 780
 Qy 781 INVKOFFNLNLEDPQEDMEVIVKCCSWASDAGARIMVAANKIPLFDIHTSGILGELIHTGHH 840
 Db 781 INVKOFFNLNLEDPQEDMEVIVKCCSWASDAGARIMVAANKIPLFDIHTSGILGELIHTGHH 840
 Qy 841 STIQVCDSPQNLAVVALSYCVELWNTDSRSKVADCRGHLSSVHGVWVSPDGSFLTS 900
 Db 841 STIQVCDSPQNLAVVALSYCVELWNTDSRSKVADCRGHLSSVHGVWVSPDGSFLTS 900
 Qy 901 SDOQIRLWETKVCNSAVMLKQEVVVVQENWVWVAVHIRELQILNGRTGQIDYLT 960
 Db 901 SDOQIRLWETKVCNSAVMLKQEVVVVQENWVWVAVHIRELQILNGRTGQIDYLT 960
 Qy 961 EQAVSCCLLSPHLQYIATGDENGAIELELVNRIPOSFOHKKTVWHIOPTADEKTLIS 1020
 Db 961 EQAVSCCLLSPHLQYIATGDENGAIELELVNRIPOSFOHKKTVWHIOPTADEKTLIS 1020
 Qy 1021 SSDDAEIQWNLQDKICIFLGHQETVKDFRLLKNSRLLSWFSFGTVKVNIIITGNKEKD 1080
 Db 1021 SSDDAEIQWNLQDKICIFLGHQETVKDFRLLKNSRLLSWFSFGTVKVNIIITGNKEKD 1080
 Qy 978 SSDDAEIQWNLQDKICIFLGHQETVKDFRLLKNSRLLSWFSFGTVKVNIIITGNKEKD 1037
 Db 978 SSDDAEIQWNLQDKICIFLGHQETVKDFRLLKNSRLLSWFSFGTVKVNIIITGNKEKD 1037

Qy 1081 FVCHQGTVLSCDISHDATKFSSTSDAKTAKIWSFDLLPLPLHELGHNGCVRCSAFSDST 1140
 Db 1038 FVCHQGTVLSCDISHDATKFSSTSDAKTAKIWSFDLLPLPLHELGHNGCVRCSAFSDST 1097
 Qy 1141 LLATGDDNGIRIRWVNSGELLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGIK 1200
 Db 1098 LLATGDDNGIRIRWVNSGELLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGIK 1157
 Qy 1201 WNVVVTGESSQTFYVTNGTNLKKIHVSPDKTVVTVDNLGLIYILQTL 1248
 Db 1158 WNVVVTGESSQTFYVTNGTNLKKIHVSPDKTVVTVDNLGLIYILQTL 1205

RESULT 2
 US-09-435-115-16
 ; Sequence 16: Application US/09435115
 ; Patent No 6346807
 ; GENERAL INFORMATION:
 ; APPLICANT: Henzel, William J
 ; TITLE OF INVENTION: APAP-1, AN ACTIVATOR OF C ASPASE-3
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
 ; STREET: 3100 No. 6346607 West Center, 90 South Seventh St
 ; CITY: Minneapolis
 ; STATE: MN
 ; COUNTRY: USA
 ; ZIP: 55402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/435,115
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/092,508
 ; FILING DATE:
 ; APPLICATION NUMBER: 60/055,258
 ; FILING DATE: 07-AUG-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kettelberger, Ph.D., Denise M
 ; REGISTRATION NUMBER: 33,924
 ; REFERENCE/DOCKET NUMBER: 11669.6USU1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 612-332-5300
 ; TELEFAX: 612-332-9081
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1205 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FRAGMENT TYPE: internal
 ; US-09-435-115-16

Query Match 96.0%; Score 6352.5; DB 3; Length 1205;
 Best Local Similarity 96.6%; Pred. No. 0;
 Matches 1205; Conservative 0; Mismatches 0; Indels 43; Gaps 1;
 Qy 1 MDKARNCLLOHREALEKDIKTSYIMDMHISDGLFTTISEEEKVNEPTQOORAAMLIKMI 60
 Db 1 MDKARNCLLOHREALEKDIKTSYIMDMHISDGLFTTISEEEKVNEPTQOORAAMLIKMI 60
 Qy 61 LKKNDSVSVFYNALLHEGYKDLAALLHDGIIPVSSSSGKDSVSGITSYVVRTVLCGGVP 120
 Db 61 LKKNDSVSVFYNALLHEGYKDLAALLHDGIIPVSSSSGKDSVSGITSYVVRTVLCGGVP 120
 Qy 121 QRPVVVTRKLVNAIOOKLSKLGKPGWVTHGMAGCKSVLAABAVRSHLSLECCFPG 180
 Db 121 QRPVVVTRKLVNAIOOKLSKLGKPGWVTHGMAGCKSVLAABAVRSHLSLECCFPG 180

Db 121 QRPVVFVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCKSVLAEEAVRDSHLLGECFP 180
181 GVHWSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEAKDRLRLMLRKHPRL 240
181 GVHWSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEAKDRLRLMLRKHPRL 240
241 ILDDVWDSWLVKAFDSQOQLITTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFN 300
241 ILDDVWDSWLVKAFDSQOQLITTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFN 300
301 MKKADLPQOASIIKECKGSPLVSLIGALLRDPFNWEYVLKOLQKQFKRIRKSSVD 360
301 MKKADLPQOASIIKECKGSPLVSLIGALLRDPFNWEYVLKOLQKQFKRIRKSSVD 360
361 YEALDEAMSIISVEMLRDIDKYITDLSILQKDVKVPKVLCTILMDMETEEVEDILQEFVN 420
361 YEALDEAMSIISVEMLRDIDKYITDLSILQKDVKVPKVLCTILMDMETEEVEDILQEFVN 420
421 KSLFLCDRNGKSFRIYLDLQVDFUTENKCSQOLHKKIITQFORHYQPHTLSPDQSDC 480
421 KSLFLCDRNGKSFRIYLDLQVDFUTENKCSQOLHKKIITQFORHYQPHTLSPDQSDC 480
481 MYWYNFLAYHVASAKWHKELCALMFLDWIKAKTELVGPAHLIHEFVEYRHLIDKCAV 540
481 MYWYNFLAYHVASAKWHKELCALMFLDWIKAKTELVGPAHLIHEFVEYRHLIDKCAV 540
541 SENFOEFLSLNGHLGRFPFNIVOLGCEPETSEVYQOAKLOAKQEVNDGMLYLEWINK 600
541 SENFOEFLSLNGHLGRFPFNIVOLGCEPETSEVYQOAKLOAKQEVNDGMLYLEWINK 600
601 KNIITNLSRLVVRPHPTDAVYHACFSDGQRIASCADKTLQVPKABTGEKLEIKAHEDV 660
601 KNIITNLSRLVVRPHPTDAVYHACFSDGQRIASCADKTLQVPKABTGEKLEIKAHEDV 660
661 LCCAFSTDRTATCSVDKVKIWNMTGELVHTYDEHSEOVNCHFTNSSHLLLATGS 720
661 LCCAFSTDRTATCSVDKVKIWNMTGELVHTYDEHSEOVNCHFTNSSHLLLATGS 720
721 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRSPDDKLLASCADGTLKLDWATSANERKS 780
721 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRSPDDKLLASCADGTLKLDWATSANERKS 780
781 INVQOFFLNLDEPQEDMEVIVKCCSWSDAGARIMVAANKIPLFDIHTSGLLGEIHTGHH 840
781 INVQOFFLNLDEPQEDMEVIVKCCSWSDAGARIMVAANKIPLFDIHTSGLLGEIHTGHH 840
841 STIQYCDSPQNHVALVAVLSQYCVLWNTDSRSKVADCRGHLWSVHGMFSPDGSSFLTS 900
841 STIQYCDSPQNHVALVAVLSQYCVLWNTDSRSKVADCRGHLWSVHGMFSPDGSSFLTS 900
857 LWNTRSRKSVADCRGHLWSVHGMFSPDGSSFLTS 857
901 SDDQTLRLWETKVKCKNSAVMLKQEVVDVFOENVMVLAVDHIRLQLINGRTGQIDYLT 960
901 SDDQTLRLWETKVKCKNSAVMLKQEVVDVFOENVMVLAVDHIRLQLINGRTGQIDYLT 960
858 SDDQTLRLWETKVKCKNSAVMLKQEVVDVFOENVMVLAVDHIRLQLINGRTGQIDYLT 917
961 EAQVSCCLSPHLOVIAFGDENGATEILELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 1020
961 EAQVSCCLSPHLOVIAFGDENGATEILELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 977
1021 SSDDAEIOVWNLQDKCIFLGHQOBTVDKDFLLKNSRLLSWFSFGTVKVMNIIITGNKEK 1080
978 SSDDAEIOVWNLQDKCIFLGHQOBTVDKDFLLKNSRLLSWFSFGTVKVMNIIITGNKEK 1037
1081 FVCHGTQVLSCDISHDATKFSSTADKATWSFDLLLPLHELGRHNCVCSAFSVDST 1140
1038 FVCHGTQVLSCDISHDATKFSSTADKATWSFDLLLPLHELGRHNCVCSAFSVDST 1097
1141 LLATGDDNGEIRIWNVSGELLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1200
1098 LLATGDDNGEIRIWNVSGELLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1157
1201 WNWVVTGSSQTFYTGNGNLKKIHVSPDPKTYVTVDNLGILYILQTL 1248

Db 1158 WNWVVTGSSQTFYTGNGNLKKIHVSPDPKTYVTVDNLGILYILQTL 1205
RESULT 3
US-09-092-508-2
; Sequence 2, Application US/09092508
; Patent No. 6291643
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6291643west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/092,508
; FILING DATE: 05-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,807
; FILING DATE: 05-JUN-1997
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-09-092-508-2
Query Match 94.9%; Score 6282; DB 3; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;
Qy 1 MDAKARNCLLOHRALEKDKITSYIMDMHISDGFLLTISEEKKVNEPTQOQRAAMLIKMI 60
Db 1 MDAKARNCLLOHRALEKDKITSYIMDMHISDGFLLTISEEKKVNEPTQOQRAAMLIKMI 60
Qy 61 LKKNDSVVSFYNALLHEGYKDLAALLHDGIPVVSSSSGKDSVSGITSYVTVLCEGVP 120
Db 61 LKKNDSVVSFYNALLHEGYKDLAALLHDGIPVVSSSSGKDSVSGITSYVTVLCEGVP 109
Qy 121 QRPVVFVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCKSVLAEEAVRDSHLLGECFP 180
Db 110 QRPVVFVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCKSVLAEEAVRDSHLLGECFP 169
Qy 181 GVHWSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEAKDRLRLMLRKHPRL 240
Db 170 GVHWSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEAKDRLRLMLRKHPRL 229
Qy 241 ILDDVWDSWLVKAFDSQOQLITTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFN 300
Db 230 ILDDVWDSWLVKAFDSQOQLITTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFN 289

```

Qy 301 MKKADLPEQAHSIIKECKGSPVWSLIGALLRFPNRYEYIKOLQNKQPKRIRKSSSYD 360
Db 290 MKKADLPEQAHSIIKECKGSPVWSLIGALLRFPNRYEYIKOLQNKQPKRIRKSSSYD 349
Qy 361 YEALDEAMSISVEMLRREDIKYITDLSILQKQVVKPTKVLCLMDMETEVEEDILQEFVN 420
Db 350 YEALDEAMSISVEMLRREDIKYITDLSILQKQVVKPTKVLCLMDMETEVEEDILQEFVN 409
Qy 421 KSLFLCDRNGKFRYYLHDLQVDFTEKNSQQLDHLKKIITQFORVHOPHTLSPQEDC 480
Db 410 KSLFLCDRNGKFRYYLHDLQVDFTEKNSQQLDHLKKIITQFORVHOPHTLSPQEDC 469
Qy 481 MYWYNFLAYHMASAKWHELCALMFSLDWKATELVGPAHLIHFVEYRHLDEKCAV 540
Db 470 MYWYNFLAYHMASAKWHELCALMFSLDWKATELVGPAHLIHFVEYRHLDEKCAV 529
Qy 541 SENFQFLSLNGHLGROPPFNIVOLGLCEPTESEVYQAKLOAKQOEVDNGMLYLEWINK 600
Db 530 SENFQFLSLNGHLGROPPFNIVOLGLCEPTESEVYQAKLOAKQOEVDNGMLYLEWINK 589
Qy 601 KNTNLURLVVRPHDTAVYHACFSEDQRIASCADKTLQVPKAPTGEKLEIKAHEDEV 660
Db 590 KNTNLURLVVRPHDTAVYHACFSEDQRIASCADKTLQVPKAPTGEKLEIKAHEDEV 649
Qy 661 LCCAFSTDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLATGS 720
Db 650 LCCAFSTDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLATGS 709
Qy 721 SDCFLKLDLNOKECRNFMGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
Db 710 SDCFLKLDLNOKECRNFMGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 769
Qy 781 INVQKQFLNLEDPQEDMEVIVKCCSWSADGARIMVAANKIPLFDIHTSGLLGEIHTGHH 840
Db 770 INVQKQFLNLEDPQEDMEVIVKCCSWSADGARIMVAANKIPLFDIHTSGLLGEIHTGHH 811
Qy 841 STIOYCDPSPQNHAAVVALSYQVELWNTDSTRSKVADCRGHLSSWVHGVMFSPDGSFLTS 900
Db 812 STIOYCDPSPQNHAAVVALSYQVELWNTDSTRSKVADCRGHLSSWVHGVMFSPDGSFLTS 846
Qy 901 SDDOTIRLWETKVKCKSAVMLKQEDVVPQENWVLAHRIHRLQOLINGRTGQIDYLT 960
Db 847 SDDOTIRLWETKVKCKSAVMLKQEDVVPQENWVLAHRIHRLQOLINGRTGQIDYLT 906
Qy 961 EAQVSCCLSPHLQYIAPGDENGAEIIELVNRRIPQSRFQHKTVWHIOFTADEKTLIS 1020
Db 907 EAQVSCCLSPHLQYIAPGDENGAEIIELVNRRIPQSRFQHKTVWHIOFTADEKTLIS 966
Qy 1021 SSDDAEIQVNNWQDKCIFLGHQETVQDFRLLKNSRLLSWSFDGTVKVWNIITGNKEKD 1080
Db 967 SSDDAEIQVNNWQDKCIFLGHQETVQDFRLLKNSRLLSWSFDGTVKVWNIITGNKEKD 1026
Qy 1081 FVCHQGTVLSCDISHDATKFSSTASADTKIWSFDLPLPLHELGHNGCVRCSAFVSVDST 1140
Db 1027 FVCHQGTVLSCDISHDATKFSSTASADTKIWSFDLPLPLHELGHNGCVRCSAFVSVDST 1086
Qy 1141 LLATGDNGEIRIWNVNGELLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1200
Db 1087 LLATGDNGEIRIWNVNGELLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1146
Qy 1201 WNNVVTGESSQFTYNGNLKKHVSPDKFTYVVDNLGILYIQLTLE 1248
Db 1147 WNNVVTGESSQFTYNGNLKKHVSPDKFTYVVDNLGILYIQLTLE 1194

```

RESULT 4

```

US-09-435-115-2
; Sequence 2, Application US/09435115
; Patent No. 6346607
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: AFAF-1, AN ACTIVATOR OF C ASPASE-3

```

```

; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6346607 West Center, 90 South Seventh St.
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-Seq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/435,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/092,508
; FILING DATE:
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-09-435-115-2

```

```

Query Match 94.9%; Score 6282; DB 3; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

```

```

Qy 1 MDKARNCLLOHREALEKDKITSYIMDHIMISDGLTISEEEKVNRNEPTQOORAMLIKMI 60
Db 1 MDKARNCLLOHREALEKDKITSYIMDHIMISDGLTISEEEKVNRNEPTQOORAMLIKMI 60
Qy 61 LKQNDSTSVFYNNALHEGYKDLAALLHDGIPVWSSSSGKDSVSGITSYVTVLCEGVP 120
Db 61 LKQNDSTSVFYNNALHEGYKDLAALLHDGIPVWSSSSGKDSVSGITSYVTVLCEGVP 109
Qy 121 QRPVVFVTRKLVNAIQOKLSKLKGEPEGWVTIHGMAGCGKSVLAABAVRDHSLLEGCPFG 180
Db 110 QRPVVFVTRKLVNAIQOKLSKLKGEPEGWVTIHGMAGCGKSVLAABAVRDHSLLEGCPFG 169
Qy 181 GVHWVSVGKQDKSGLMKLQNLCTRLDQDESFSORLPLNIEEAKDRLRILMLRKHPRSL 240
Db 170 GVHWVSVGKQDKSGLMKLQNLCTRLDQDESFSORLPLNIEEAKDRLRILMLRKHPRSL 229
Qy 241 ILDDWDSWVLKAFDSCQCIILTTTRDKSVTDVSMGPKYVVPVSSSLCKEKGLEILSLFVN 300
Db 230 ILDDWDSWVLKAFDSCQCIILTTTRDKSVTDVSMGPKYVVPVSSSLCKEKGLEILSLFVN 289
Qy 301 MKKADLPEQAHSIIKECKGSPVWSLIGALLRFPNRYEYIKOLQNKQPKRIRKSSSYD 360
Db 290 MKKADLPEQAHSIIKECKGSPVWSLIGALLRFPNRYEYIKOLQNKQPKRIRKSSSYD 349
Qy 361 YEALDEAMSISVEMLRREDIKYITDLSILQKQVVKPTKVLCLMDMETEVEEDILQEFVN 420
Db 350 YEALDEAMSISVEMLRREDIKYITDLSILQKQVVKPTKVLCLMDMETEVEEDILQEFVN 409
Qy 421 KSLFLCDRNGKFRYYLHDLQVDFTEKNSQQLDHLKKIITQFORVHOPHTLSPQEDC 480

```

Db 410 KSLFLCDNRNGSFYIYLDQVDFTEKNCSQLDLHKKIITQFORIHOPHTLSPDQDC 469
Qy 481 MYWYNFLAYHMASAKMHELKALMFLSDWIKAKTELAVGPAHLIHEFVEYRHLDEKCAV 540
Db 470 MYWYNFLAYHMASAKMHELKALMFLSDWIKAKTELAVGPAHLIHEFVEYRHLDEKCAV 529
Qy 541 SENFOEFLSLNGHLLGRQPPFNIVOLGLCEPETSSEYVQAKLQAKQEVNDGMLYLEWINK 600
Db 530 SENFOEFLSLNGHLLGRQPPFNIVOLGLCEPETSSEYVQAKLQAKQEVNDGMLYLEWINK 589
Qy 601 KNIITNLSRLVVRPHTDAVTHACFSDGQRIASCADKTLQVFKAETGKLEIKAHEDV 660
Db 590 KNIITNLSRLVVRPHTDAVTHACFSDGQRIASCADKTLQVFKAETGKLEIKAHEDV 649
Qy 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS 720
Db 650 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS 709
Qy 721 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
Db 710 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 769
Qy 781 INVKQFFLNLEDPQEDMEVIVKCCSWSDAGARIMVAANKIFLFDIHTSGLLGEIHTGHH 840
Db 770 INVKQFFLNLEDPQEDMEVIVKCCSWSDAGARIMVAANKIFLFDIHTSGLLGEIHTGHH 811
Qy 841 STIOYCFSPQNHVALVALSQVELWNTDSRSKVDRCRHLGHSVGHVGMFSPDGSFPTS 900
Db 812 -----LWNTDSRSKVDRCRHLGHSVGHVGMFSPDGSFPTS 846
Qy 901 SDDQITRLWETKVKCNKSAVMLKQEVNVFOENEVVLAVDHIRLQOLINGRTGQIDYLT 960
Db 847 SDDQITRLWETKVKCNKSAVMLKQEVNVFOENEVVLAVDHIRLQOLINGRTGQIDYLT 906
Qy 961 EAOVSCCLSPHLQYIAFGDENGAIIELELVNRRIFQSRFOHKTVMWHIQTADKTLIS 1020
Db 907 EAOVSCCLSPHLQYIAFGDENGAIIELELVNRRIFQSRFOHKTVMWHIQTADKTLIS 966
Qy 1021 SSDDAEIQVWNNQDKCIFLRGHQBTVDKFLKNSRLLSWSDGTVMVNNIITGNKEK 1080
Db 967 SSDDAEIQVWNNQDKCIFLRGHQBTVDKFLKNSRLLSWSDGTVMVNNIITGNKEK 1026
Qy 1081 FVCHQGTVLSCDISHDATKFSSTSDAKTIWSPDLLPLHELGHNGCVRCSAFSDST 1140
Db 1027 FVCHQGTVLSCDISHDATKFSSTSDAKTIWSPDLLPLHELGHNGCVRCSAFSDST 1086
Qy 1141 LLATGDDNGEIRIWNVSGELLHLCAPLSEGAATHGGVTDLCFSPDGKMLISAGGYIK 1200
Db 1087 LLATGDDNGEIRIWNVSGELLHLCAPLSEGAATHGGVTDLCFSPDGKMLISAGGYIK 1146
Qy 1201 WNNVVTGSSOTFYTNGTNLKKIHVSPDKTAVTVDNIGILYIQTLE 1248
Db 1147 WNNVVTGSSOTFYTNGTNLKKIHVSPDKTAVTVDNIGILYIQTLE 1194

RESULT 5
US-09-069-023-26
; Sequence 26, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 1194

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-26
Query Match 94.9%; Score 6282; DB 3; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;
Qy 1 MDKARNCLLOHREALEKDIKTSYIMDHISDGFITISEEEKVRNEPTQQQRAAMLKMI 60
Db 1 MDKARNCLLOHREALEKDIKTSYIMDHISDGFITISEEEKVRNEPTQQQRAAMLKMI 60
Qy 61 LKKNDSVTSFYNALLHGGYKDLAALLHDGIPVSSSSSGKDSVSITSVYRTVLCEGGVP 120
Db 61 LKKNDSVTSFYNALLHGGYKDLAALLHDGIPVSSSSSGKDSVSITSVYRTVLCEGGVP 109
Qy 121 QRPVVFTVRKKLVNAIQOKLSKLKGEPCGWVTHGMAGCGKSVLAEEAVRDSHLLSGCFPG 180
Db 110 QRPVVFTVRKKLVNAIQOKLSKLKGEPCGWVTHGMAGCGKSVLAEEAVRDSHLLSGCFPG 169
Qy 181 GVVHVSVGKQDKSGLLMKLNLCRLDQDESFSQRLPLNIEEAKDRLILMLRKHPSLL 240
Db 170 GVVHVSVGKQDKSGLLMKLNLCRLDQDESFSQRLPLNIEEAKDRLILMLRKHPSLL 229
Qy 241 ILDDVMSWVLKAFDSQCQIILLTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN 300
Db 230 ILDDVMSWVLKAFDSQCQIILLTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN 289
Qy 301 MKKADLPEQAHSIIKECKGSPVLSLIGALLRDPNRWEYVYLKOLNKFQKRIKSSSYD 360
Db 290 MKKADLPEQAHSIIKECKGSPVLSLIGALLRDPNRWEYVYLKOLNKFQKRIKSSSYD 349
Qy 361 YEALDEAMSISVEMLRREDIKDYTDLSILOKDVKPTKVLGILWDMETEEVEDILOEFVN 420
Db 350 YEALDEAMSISVEMLRREDIKDYTDLSILOKDVKPTKVLGILWDMETEEVEDILOEFVN 409
Qy 421 KSLFLCDNRNGSFYIYLDQVDFTEKNCSQLDLHKKIITQFORIHOPHTLSPDQDC 480
Db 410 KSLFLCDNRNGSFYIYLDQVDFTEKNCSQLDLHKKIITQFORIHOPHTLSPDQDC 469
Qy 481 MYWYNFLAYHMASAKMHELKALMFLSDWIKAKTELAVGPAHLIHEFVEYRHLDEKCAV 540
Db 470 MYWYNFLAYHMASAKMHELKALMFLSDWIKAKTELAVGPAHLIHEFVEYRHLDEKCAV 529
Qy 541 SENFOEFLSLNGHLLGRQPPFNIVOLGLCEPETSSEYVQAKLQAKQEVNDGMLYLEWINK 600
Db 530 SENFOEFLSLNGHLLGRQPPFNIVOLGLCEPETSSEYVQAKLQAKQEVNDGMLYLEWINK 589
Qy 601 KNIITNLSRLVVRPHTDAVTHACFSDGQRIASCADKTLQVFKAETGKLEIKAHEDV 660
Db 590 KNIITNLSRLVVRPHTDAVTHACFSDGQRIASCADKTLQVFKAETGKLEIKAHEDV 649
Qy 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS 720
Db 650 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS 709
Qy 721 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
Db 710 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 769
Qy 781 INVKQFFLNLEDPQEDMEVIVKCCSWSDAGARIMVAANKIFLFDIHTSGLLGEIHTGHH 840
Db 770 INVKQFFLNLEDPQEDMEVIVKCCSWSDAGARIMVAANKIFLFDIHTSGLLGEIHTGHH 811
Qy 841 STIOYCFSPQNHVALVALSQVELWNTDSRSKVDRCRHLGHSVGHVGMFSPDGSFPTS 900
Db 812 -----LWNTDSRSKVDRCRHLGHSVGHVGMFSPDGSFPTS 846
Qy 901 SDDQITRLWETKVKCNKSAVMLKQEVNVFOENEVVLAVDHIRLQOLINGRTGQIDYLT 960
Db 847 SDDQITRLWETKVKCNKSAVMLKQEVNVFOENEVVLAVDHIRLQOLINGRTGQIDYLT 906
Qy 961 EAOVSCCLSPHLQYIAFGDENGAIIELELVNRRIFQSRFOHKTVMWHIQTADKTLIS 1020

```

Db 907 EAQVSCCLSPHLQYIAFGDNGAIEIIELVNRRIFQSRFOHKTVMHIQFTADEKTLIS 966
Qy 1021 SSDASIQVNNWQDKCIFLRGHQETVKDPRLLKNSRLLSWSDGTGVKVMNIIITGNKEKD 1080
Db 967 SSDASIQVNNWQDKCIFLRGHQETVKDPRLLKNSRLLSWSDGTGVKVMNIIITGNKEKD 1026
Qy 1081 FVCHQGTVLSCDISHDATKFSSTADTKAKIWSFDLPLHLRHGNCVRCSAFSDVST 1140
Db 1027 FVCHQGTVLSCDISHDATKFSSTADTKAKIWSFDLPLHLRHGNCVRCSAFSDVST 1086
Qy 1141 LLATGDNGEIRIWNVSNGLLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGIYK 1200
Db 1087 LLATGDNGEIRIWNVSNGLLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGIYK 1146
Qy 1201 WNVVTGESSQTFYTNGLNKKIHVSPDKTYVTVDNLGILYILOTLE 1248
Db 1147 WNVVTGESSQTFYTNGLNKKIHVSPDKTYVTVDNLGILYILOTLE 1194

RESULT 6
US-09-098-310-2
; Sequence 2, Application US/09098310
; Patent No. 6403765
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emd S.
; TITLE OF INVENTION: TRUNCATED APAF-1 AND METHODS OF USE
; FILE REFERENCE: 480140.438
; CURRENT APPLICATION NUMBER: US/09/098.310
; PRIORITY FILING DATE: 1998-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1194
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-098-310-2

Query Match 94.9%; Score 6282; DB 4; Length 1194;
Best Local Similarity 95.7%; Pred No 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

Qy 1 MDKARNCLLOHREALEKDICTSYMDHMSDGLFTISEEKVNRNEPTQOORAMLIKMI 60
Db 1 MDKARNCLLOHREALEKDICTSYMDHMSDGLFTISEEKVNRNEPTQOORAMLIKMI 60
Qy 61 LKDNDSYVSFYNALHHEGYKDLAALLHDGIPVWSSSGKDSVSGITSYRTVLCRGGVP 120
Db 61 LKDNDSYVSFYNALHHEGYKDLAALLHDGIPVWSSSGKDSVSGITSYRTVLCRGGVP 109
Qy 121 QRPVVFVTRKKLVNAIQKLSKLKGPVWVTHGMAGCKSVLAEEAVRDHSLLEGCPFG 180
Db 110 QRPVVFVTRKKLVNAIQKLSKLKGPVWVTHGMAGCKSVLAEEAVRDHSLLEGCPFG 169
Qy 181 GVHWVSVGKQDGLLMLQNLCTRLDQDESFSQRLPLNIEAKDRILMLRKHPRSL 240
Db 170 GVHWVSVGKQDGLLMLQNLCTRLDQDESFSQRLPLNIEAKDRILMLRKHPRSL 229
Qy 241 ILDDVWDSWLVKAPDSQOILLTTTRDKSVTDSVMGPKVVPVSSSLGKGLBLSLFVN 300
Db 230 ILDDVWDSWLVKAPDSQOILLTTTRDKSVTDSVMGPKVVPVSSSLGKGLBLSLFVN 289
Qy 301 MKKADLPEQAHSIIECKGSPVLSLIGALLRDFNRWEYVLKOLQNKQPKRIRKSSSYD 360
Db 290 MKKADLPEQAHSIIECKGSPVLSLIGALLRDFNRWEYVLKOLQNKQPKRIRKSSSYD 349
Qy 361 YEALDEAMSIIVEMREDIKDYTDLSILQKQVPTKVLCLMDMTEVEEDILQEFVN 420
Db 350 YEALDEAMSIIVEMREDIKDYTDLSILQKQVPTKVLCLMDMTEVEEDILQEFVN 409
Qy 421 KSLLECDNRNGKSFYRLHLDQVDFTEKNCSQLDLHKKIITQFQRYHQPHLSPQDEDC 480

```

```

Db 410 KSLLECDNRNGKSFYRLHLDQVDFTEKNCSQLDLHKKIITQFQRYHQPHLSPQDEDC 469
Qy 481 MTWYNFLAYHMASAKMHELCALMFSLDWIKAKTELGVPAHLIHEFVEYRHILDEKDCAV 540
Db 470 MTWYNFLAYHMASAKMHELCALMFSLDWIKAKTELGVPAHLIHEFVEYRHILDEKDCAV 529
Qy 541 SENFOEFLSLNGLHLLGRPPNIVOLGLCEPETSIVYQOAKLOAKQEVNDGMLYLEWINK 600
Db 530 SENFOEFLSLNGLHLLGRPPNIVOLGLCEPETSIVYQOAKLOAKQEVNDGMLYLEWINK 589
Qy 601 KNITNLSLVVRPHTDAVYHACFSESDGRIASCAGADTKLOVFKAETGKLEIKAHEDV 660
Db 590 KNITNLSLVVRPHTDAVYHACFSESDGRIASCAGADTKLOVFKAETGKLEIKAHEDV 649
Qy 661 LCCARSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLIATGS 720
Db 650 LCCARSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLIATGS 709
Qy 721 SDCFLKJLMDLNQKECRNTPFGHTNSVNHCRFSPDDKLLASCADGTJLKLWDATSAHERKS 780
Db 710 SDCFLKJLMDLNQKECRNTPFGHTNSVNHCRFSPDDKLLASCADGTJLKLWDATSAHERKS 769
Qy 781 INVKQFFLNLEDPQEDMEVIVKCCSWSDAGARIMVAAKNKIFLFDIHTSGLLGEIHTGHH 840
Db 770 INVKQFFLNLEDPQEDMEVIVKCCSWSDAGARIMVAAKNKIF 811
Qy 841 STIQCYDFSPQNHLAVVALLSYCYVELWNTDSRSKVDACRGLHSVWVGWVMSPDGSSFLTS 900
Db 812 STIQCYDFSPQNHLAVVALLSYCYVELWNTDSRSKVDACRGLHSVWVGWVMSPDGSSFLTS 846
Qy 901 SDDQITRLWETKVKCKNSAVMLKQEVVDVVOFQENEVWVLAVDHRRRLQINGRTGQIDYLT 960
Db 847 SDDQITRLWETKVKCKNSAVMLKQEVVDVVOFQENEVWVLAVDHRRRLQINGRTGQIDYLT 906
Qy 961 EAQVSCCLSPHLQYIAFGDNGAIEIIELVNRRIFQSRFOHKTVMHIQFTADEKTLIS 1020
Db 907 EAQVSCCLSPHLQYIAFGDNGAIEIIELVNRRIFQSRFOHKTVMHIQFTADEKTLIS 966
Qy 1021 SSDASIQVNNWQDKCIFLRGHQETVKDPRLLKNSRLLSWSDGTGVKVMNIIITGNKEKD 1080
Db 967 SSDASIQVNNWQDKCIFLRGHQETVKDPRLLKNSRLLSWSDGTGVKVMNIIITGNKEKD 1026
Qy 1081 FVCHQGTVLSCDISHDATKFSSTADTKAKIWSFDLPLHLRHGNCVRCSAFSDVST 1140
Db 1027 FVCHQGTVLSCDISHDATKFSSTADTKAKIWSFDLPLHLRHGNCVRCSAFSDVST 1086
Qy 1141 LLATGDNGEIRIWNVSNGLLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGIYK 1200
Db 1087 LLATGDNGEIRIWNVSNGLLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGIYK 1146
Qy 1201 WNVVTGESSQTFYTNGLNKKIHVSPDKTYVTVDNLGILYILOTLE 1248
Db 1147 WNVVTGESSQTFYTNGLNKKIHVSPDKTYVTVDNLGILYILOTLE 1194

RESULT 7
US-09-538-092-825
; Sequence 825, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538.092
; PRIORITY FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqformatter Version 0.9
; SEQ ID NO 825

```

LENGTH: 1194
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)-(0)
; OTHER INFORMATION: Polypeptide Accession Number O14727
US-09-538-092-825

Query Match 94.9%; Score 6282; DB 4; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

Qy 1 MDAKARNCLLQHRBALEKDIKTSYIMDMHISDGFLTISEEBKVRNEPTQQQRAAMLIKMI 60
Db 1 MDAKARNCLLQHRBALEKDIKTSYIMDMHISDGFLTISEEBKVRNEPTQQQRAAMLIKMI 60

Qy 61 LKKNDSSVVSFNALLHEGYKDLAALLHDGIPVVSSSGKDSVSGITSYVTVLCEGVP 120
Db 61 LKKNDSSVVSFNALLHEGYKDLAALLHDGIPVVSSSGKDSVSGITSYVTVLCEGVP 109

Qy 121 QRPVVFTRKKLVNAIQKLSKLGEPGWVTHGMAGCGKSVLAEEAVRDHSLLEGCPFG 180
Db 110 QRPVVFTRKKLVNAIQKLSKLGEPGWVTHGMAGCGKSVLAEEAVRDHSLLEGCPFG 169

Qy 181 GVHVSVGQKSGLLMKLQNLCTRLDQESFSQRLPLNIEEAKDRILMLRKHPRSL 240
Db 170 GVHVSVGQKSGLLMKLQNLCTRLDQESFSQRLPLNIEEAKDRILMLRKHPRSL 229

Qy 241 ILDDVMSWLKAFDSQCIILTRDKSVTVSVMGPKVVPVSSSLGKEGLEILSLFVN 300
Db 230 ILDDVMSWLKAFDSQCIILTRDKSVTVSVMGPKVVPVSSSLGKEGLEILSLFVN 289

Qy 301 MKKADLPQAHSIKECKGSLVSLIGALLRDPNREYVYLKOLNKPFRIRKSSYD 360
Db 290 MKKADLPQAHSIKECKGSLVSLIGALLRDPNREYVYLKOLNKPFRIRKSSYD 349

Qy 361 YEALDEAMSIIVEMLRDIDKYTDLSTLQKDVKVPYTKVLCILMDMETEEVEDILOEPVN 420
Db 350 YEALDEAMSIIVEMLRDIDKYTDLSTLQKDVKVPYTKVLCILMDMETEEVEDILOEPVN 409

Qy 421 KSLIFCDRNGSKFRYLHDLQVDFLTKNCSQQLQDLHKKIITQORHQPHLSPDQEDC 480
Db 410 KSLIFCDRNGSKFRYLHDLQVDFLTKNCSQQLQDLHKKIITQORHQPHLSPDQEDC 469

Qy 481 MYWTFNFLAYHMASAKMHELCALMFLSDWIKAKTELVPAPLIIHEFVEYRHLIDKCAV 540
Db 470 MYWTFNFLAYHMASAKMHELCALMFLSDWIKAKTELVPAPLIIHEFVEYRHLIDKCAV 529

Qy 541 SENFQFLSLNGLHLLGRQFPFNIIVQLGLCEPETSEVYQQAQKQAEVDNGLYLEWINK 600
Db 530 SENFQFLSLNGLHLLGRQFPFNIIVQLGLCEPETSEVYQQAQKQAEVDNGLYLEWINK 589

Qy 601 KNIITNLSRLVVRPHTDAVYHACFSDGQRIASCADKTLQVFKAEKGKLEIKAEHDEV 660
Db 590 KNIITNLSRLVVRPHTDAVYHACFSDGQRIASCADKTLQVFKAEKGKLEIKAEHDEV 649

Qy 661 LCCAFSTDRTIATCSVDKVKIWNMTGELVHTYDHSQVNCCHFTNSHLLLATGS 720
Db 650 LCCAFSTDRTIATCSVDKVKIWNMTGELVHTYDHSQVNCCHFTNSHLLLATGS 709

Qy 721 SDCLFLKLDLWLNQKCRNTFMGTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 780
Db 710 SDCLFLKLDLWLNQKCRNTFMGTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 769

Qy 781 INVQKFFLNLEDPOEDMEVIVKCCSWSDAGRIWVAANKIPLFDIHTSGLLGIHTGHH 840
Db 770 INVQKFFLNLEDPOEDMEVIVKCCSWSDAGRIWVAANKIPLFDIHTSGLLGIHTGHH 811

Qy 841 STIQYCDSPQNHILAVVALSYCVELWNTDSTRSKVADCRGHLVHGVWMPSPDGSSFLT 900
Db 812 LMTNDRSKVADCRGHLVHGVWMPSPDGSSFLT 846

Qy 901 SDDOTIRLWETKVKCKNSAVMLKQEVVVFOENEVWMLAVDHIIRRLQILINGRTGQIDYLT 960
Db 847 SDDOTIRLWETKVKCKNSAVMLKQEVVVFOENEVWMLAVDHIIRRLQILINGRTGQIDYLT 906

Qy 961 EAOVSCCLSPHLOVIAFGDENGAIIELELVNRIFQSRFOHKKTWHIQFTADEKTLIS 1020
Db 907 EAOVSCCLSPHLOVIAFGDENGAIIELELVNRIFQSRFOHKKTWHIQFTADEKTLIS 966

Qy 1021 SSDDAEIQVNNWQDKCIFLRGHQETVKDFRLLKNRLLSWSFDGTVKVMNIITGNKEKD 1080
Db 967 SSDDAEIQVNNWQDKCIFLRGHQETVKDFRLLKNRLLSWSFDGTVKVMNIITGNKEKD 1026

Qy 1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLPLHLRHGNCVRCASFVSDST 1140
Db 1027 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLPLHLRHGNCVRCASFVSDST 1086

Qy 1141 LLATGDDNGEIRIWNVNGELLHLICAPLSEBGAATHGGWVTDLCFSPDGKMLISAGGYIK 1200
Db 1087 LLATGDDNGEIRIWNVNGELLHLICAPLSEBGAATHGGWVTDLCFSPDGKMLISAGGYIK 1146

Qy 1201 WNVVVTGESSQTFYNTGNTLNKKIHVSPDKTYVTVDNLGILYILQTL 1248
Db 1147 WNVVVTGESSQTFYNTGNTLNKKIHVSPDKTYVTVDNLGILYILQTL 1194

RESULT 8
US-08-751-189-3
; Sequence 3, Application US/08751189
; Patent No. 5919656
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 5919656el Genes Encoding Telomerase Protein
; TITLE OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08751,189
; APPLICATION NUMBER: US/08751,189
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleksi, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2627 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-751-189-3

Query Match 6.2%; Score 408; DB 2; Length 2627;
Best Local Similarity 21.0%; Pred. No. 4.1e-27;
Matches 296; Conservative 170; Mismatches 492; Indels 454; Gaps 58;

Qy 97 SSGKDSVSGITSYVTVLCE-----GGVQRPVVVTRKKLVNAIQKLSK-- 142
Db 1083 AAGRPYVGGLEEFQVLQDVWNNIQLYLPQALLEGQ-VSIPDDDLVQVTFQQLQKPP 1141

Qy 143 -----LKGEPPGWVTHGMAGCGKSVLAEEAVRDHSLLEGCPFGV--- 182

```

Db 1142 SPARPLQLDVTQXLMPLPHGLSLV---GSGGKGTAFIASLV---SALQA---PDGAKVA 1194
Qy 183 -----HWVGVGQDKSGLLMKQLNCTRLDQDSEFSQRLPLNIEAKDRILMLRKHPH 237
Db 1195 XLVFFHF-SGAPDQGLALTLLRLCTYLRGQLKEGALFSTYSLVWELQORLLPKSAE 1253
Qy 238 SL-----LILD-----DWDWSVLKAFDSQCOILLTRDKSDVTSVWVGPKYVV 280
Db 1254 SLHPGQTVLIIDGADRLVDQGLISDWTPKPLPCRVHLVSVS---SDAGLG-----1304
Qy 281 PVESLGEKEGLEILSLFVNNKKADLPQEAHSIIKECKSPVLVSLIGALIRLPPNRWEY 340
Db 1305 ---ETLEQSQGAHVLAAL-----GPLEASARALVRE-----ELAL 1336
Qy 341 YLKOLQNKQFKR-----IRKSSDYDEALDEAMSISVEMLRDIDKYVTDLSLOKDVK 394
Db 1337 YGKRLESPFNQMRLLLVKRESGRPLVRLVTDHLRLFTLYEQVSRRLTL-----PAT 1391
Qy 395 VPTKVLCLMDMETEEVEDILOEP-----VNSKLLFCDR-----428
Db 1392 VPLLQHLILSLKEHGDVLPQALTALEVTRESGLTVQDLHGVLVSWRTLPGKTKSWEA 1451
Qy 429 --NGKSFYX-----LHDLOVDPLTEKNCOSLODLHKLIITOFORVHOPHTLSP 475
Db 1452 VAAGNSGDPYMGFPACLVQSLRSLGEGPLRPGARCLDPGLPRTAAKCYKCR---P 1508
Qy 476 DQEDCMY-----W-----YNFLAYHMASAKMKELCALMFSLDWI 510
Db 1509 GLEDTAHILIAQLMKTCDASGTFRSCPEPMALGDPHYHLLQSGNRLGSLKFTLNHLV 1568
Qy 511 KAKTELVGPAHLIHEFYV-----RHILDEKDCVNSENFOEFLSNGHLRQRPFPNI 563
Db 1569 AHLEGLVSLRLELAHALYASVPKEQKLFEADVAV---FRFLRQOASLTQO---YPRL 1623
Qy 564 V-QLGCEPETSEVYQQAQLOAK-----QEVNDGMVLEWINK---XNITNLS-RLVVRP 613
Db 1624 LPQQAANQDPLSPCHQASLSLRWHLOHT-----LRLANKPTMKNQOSSLSLAVS 1677
Qy 614 HTDAVTHAFCSDDQRTASCAADKTLQVKAETGEKLEIKAHEDVLCCAFSTDDRFIA 673
Db 1678 SPTAV---AFSTNGORAAVGTANGTVYLLDLRTWQEKSVVSGCDGISACLFSDDTFL 1734
Qy 674 TCSVDKVKIWNMTGELVHTYDHSBQVNCCHFTNSHHLLLATGSSDCFLKLWDINQK 733
Db 1735 T-----AFDGLLELWDL-QH 1748
Qy 734 ECRNTMF-GHTNSVNHCRFSPDDKLLASCSADGTLKMD-----ATSANERKSINVKOF 786
Db 1749 GCKVLQTKAHQVQITGCLSPDCRLLATVCLGCLKMDTVRGQLAFQHTYPKSLNCVAF 1808
Qy 787 FLNLEDPQEDMEVIKCCSWs-----ADGARIMVAANKIFLFDIHTSGLLGEIHTGH 839
Db 1809 -----HPEQO---VIATGSMAGSISFPQVDGLKVTK-----DLGAPG-----1842
Qy 840 HSTIOYCDPSQNHVLAVVALSOYVELWNTDSKSVADCRHLSVHGVWFSPDGSSFLT 899
Db 1843 -ASIRTLAFNPGVAVGRLLDSNVELMAWREGARLAAPFAHGHVAAALFLHAGCOLIT 1901
Qy 900 SDDQTLRLWE-----TKVCXSNVLMKQEVVDVFOENVMVNLAVDHIRLQ 948
Db 1902 AGEDGKGVQWSSGLRPRHGLSLSISPALSVALSPDGRV-----AVGRADGIRIYKI 1956
Qy 949 INGRGTQIDYLTEAQVSCCC-LSPHLYQYAFDENGALIELEL-----VNNRIFQSPQHK 1003
Db 1957 SSGSQGAQOQALDVAVSALAWLSP--KVLVSGAEDGSLQGWALKECSLQSLMLLSRPQ--2012
Qy 1004 KTWMIHQFTADKTLIISDDAEIQVNNMQL-----DKC-IFLRHQETV-----1047
Db 2013 KPV--LGLATSOELLASASEFTVQLWPRQLLTPRPHKAEDPCGTETLGRHGVSCCSFS 2070
Qy 1048 -----KDFRLL-----KNSRLLSWSPDGTV 1067

```

```

Db 2071 TDGSLATGGRDRRELLCWDVRRTPKTPVLIIHSPFACHRDWVTGCANTKDNLLISCSSDGSV 2130
Qy 1068 KVMNIITNKEKDFVCHQGTVL-----SCDI 1093
Db 2131 GLWDPESQRLGQPLGHQSASVAVAEEHVVSVRDGLTKVWDHOGVELTSPAHSGPI 2190
Qy 1094 SHDATKFSSTSA-----DKTAKIWSFDLLLPLHLRHNHCVCRCASFVS 1139
Db 2191 SHCAAMEPRAAQPGQSGSELVVTGLDGRATRLMHPLLVCQHTTLHGSHGVPRAAAVSETS 2250
Qy 1140 TLLATGDDNGEIRIWNYSNGELIHLCAPLSEGAATHGWWVTDLCFSPDGKMLIS---AG 1196
Db 2251 GLMLTASDGSVRLWQVPK-EADDTTCIPRSSAA-----VTAVAWAPDGSMAVSGNQAG 2302
Qy 1197 GYIKWN---VVTGE-----SSQTFY 1214
Db 2303 ELILWQEKAVATAQAPGHIGALINSSAHTFF 2334

RESULT 9
US-09-060-836-3
; Sequence 3, Application US/09060836
; Patent No. 5981707
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 5981707el Genes Encoding Telomerase Protein
; TITLE OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,836
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/751,189
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2627 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-060-836-3

Query Match 6.2%; Score 408; DB 2; Length 2627;
Best Local Similarity 21.0%; Pred. No. 4.1e-27;
Matches 296; Conservative 170; Mismatches 492; Indels 454; Gaps 58;
Qy 97 SSGKDSVSGTISVTVLCE-----GGVOPRPVTVTRKKLVNATQOKLSK-- 142
Db 1083 AAGRPPYVGGLEFEGQLVQDVNMVNTQKLYLQFGALLSQP-VSIPDDDLVQATFQQLQKPP 1141
Qy 143 -----LKGEFGWVTTHGMAGCKSVLAEEAVRDHSLLEGCFFPGGV--- 182
Db 1142 SPARPLQLDVTQXLMPLPHGLSLV---GSGGKGTAFIASLV---SALQA---PDGAKVA 1194

```


183 QY -----HWVSVGKQKSGILMKLONLCTRLDQESFQRLPLNIEBAKDRILMLRKHPR 237
1195 Db XLVPFHP-SGARPQGLALTLLRLCTYLRGOLKEPGALPSTYRSLVWELQORLLPKSAE 1253
238 QY -----LILD-----DWDMSWLKAFDSQOQIILLTRDKSVTDSVMPKVVV 280
1254 Db SLHPGQTOVLIIDGADRLVDQNGQLISDWIIPKLPKRCVHLVLSV-----SDAGLG-----1304
281 QY PVEISLKGKGLLEILSLFVNMKKADLPQAHSHIIECKGSPILVSLIGALLRDPFNWREY 340
1305 Db -----ETLEQSQAHVLA-----GPLEASARARLVRE-----ELAL 1336
341 QY YLKQONKQFPR-----IRKSSSYDYALDEAMSIISVEMLREDIKYDTLILQKDVK 394
1337 Db YGKRLESPFNQMLRLLVRESGRPLYRLVTDHLRLFTLYEQVSRRLTL-----PAT 1391
395 QY VPTKVLCTILDMETREVEDILQEF-----VNKSLFLCDR-----428
1392 Db VPLLLOHILSLTEKEHGPDPVLPQALTALVTRSGLTVDQLHGVLSVWRTLPGTKYSWEEA 1451
429 QY -----NGKSPRY-----LHDLDQVDFLTERKNSQLQDLHKIITQFRIYHOPHTLSP 475
1452 Db VAAGNSGDPYPMGPACVLQSLRSLGSGPLRPGARCLPDGPLRTAAKCYGKR---P 1508
476 QY DOEDCMY-----W-----YNFLAYHMASAKWHKELCALMFSLDWI 510
1509 Db GLEDTAHLIAAQLWKTCDDASGTFTRSCPPEALGDLFPYHLLQSGNRGLLSKFTLNHV 1568
511 QY KAKTELGVPAHLIHEFVY-----RHILDEKCAVSENFQEPFLSLNGLHLLGRPPNI 563
1569 Db AAHLLEGLVSLLENAHALYASVPKEQKLEADVAV---FRTFLRQOASILSQ---YPRL 1623
564 QY V-QLGLCBPETSSEVYQAKLOAK-----QEVDMGMLYLEWINK-----KNITNLS-RLVVRP 613
1624 Db LPQQAANQLDPLCHQASLSLRWHLOHT-----LRWLKPRTMKNQOQSSLSLAVSS 1677
614 QY HTDVAHYACFEDGRIASCADKTLQVFKAEKLEIKAEHDEVLCACAFSTDDRIA 673
1678 Db SPTAV---AFSTNGORAAVGTANGTVLLDRTQOEKSVSVSGCDGISACFLFSDDTLFL 1734
674 QY TCSVDKVKYIWNMTGELVHYDEHSEQVNCCHFTNSSHLLLATGSSDCFLKWLNDLQK 733
1735 Db T-----AFDGLLELWDL-QH 1748
734 QY ECRNTMF-GHTNSVNHCRFSPDDKLLASCADGTLKLD-----ATSANERKSNVVKQF 786
1749 Db GCRVLQTKAHQYQITGCCLPDCLLATVCLGGCLKLDVTRGQLAFQHTYTPKSLNCVAF 1808
787 QY FLNLEDPQEDMEVIVKCCWS-----ADGARINVAANKIFLFDIHTSGLLGEIHTGH 839
1809 Db -----HPEGQ---VIATGSWAGSISFFQVDGLKVTK-----DLGAPG-----1842
840 QY HSTIQYCDPSQNHVALVALSOYCVELWNTDSRKVADCRGHLVSVHGVMSFSPDSSFLT 899
1843 Db ASIRTLAFNVPGVAVVGRLDMSVLMWREGARLAAPPAHGFVAAALFLHAGCQLLT 1901
900 QY SSDDQITRLWE-----TKKVCKNSAMLKQEVDDVVVFQENVMVLAVDHIRRLQL 948
1902 Db AGEDGKVQVSGSLRPRGHLSLSLSPALSVALSPDGRV-----AVGYRADGIRIYKI 1956
949 QY INGRGTQIDYITEAQVSCC-LSPHLQYIAFGDENGATIELEL-----VNNRIFOSRFOHK 1003
1957 Db SSGSQAGQALDVAVALWALSP--KVLVSGAEDGSLQGWALKECSLQSLWLALSFRQ--2012
1004 QY KTVWHIQTAEKTLISSDDAEQVNWOL-----DKC-IFLRCHQETV-----1047
2013 Db KPV--LGLATSOELASASEFTQLWPRQLLTRPHKAEDPPCGTELHGHGPGVSCCSFS 2070
1048 QY -----KDFRLL-----KNSRLLSWSFDGTV 1067
2071 Db TDGSLATGGDRSLLDVDRTPKTPVLHSPACHRDVWTGCATWKDNLILSCSDGSV 2130
1068 QY KWNIIITGNKEKDFVCHQGTVL-----SCDI 1093

2131 Db GLWDPESQRLGQFLGHQSAVSAVAEEHVSVSRDGTCLKWDHQGVVELTSIPAHSGP 2190
1094 QY SHDATKFSSTSA-----DKTAKIWSFDLLPLHLHELGRHNGCVRCSAFSVD 1139
2191 Db SHCAAMEPRAAGQPGSELLVVTVGLDGTATRLMHPLLVCQHTLLGHSGPVRAAAVSETS 2250
1140 QY TLLATGDNGBIRIWNVNSGELLHLHLCAPLSEGAATHGWTDLCLFSPDGKMLI9---AG 1196
2251 Db GLMLTASGDGSRVLQVPK-EADDTCTIPRSSAA-----VTAVAWAPDGSMAVSGNOAG 2302
1197 QY GYIKWVN---VVTGE-----SSQTEY 1214
2303 Db ELILWQEKAVATAQAPGHIGALIWSSAHTFF 2334

RESULT 10
US-09-184-445-3
; Sequence 3, Application US/09184445
; Patent No. 6174703
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 6174703el Genes Encoding Telomerase Protein
; TITLE OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/184,445
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/751,189
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2627 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-184-445-3

Query Match 6.2%; Score 408; DB 3; Length 2627;
Best Local Similarity 21.0%; Pred. No. 4.1e-27;
Matches 296; Conservative 170; Mismatches 492; Indels 454; Gaps 58;

97 SSGKSDSVSGTISYVTVLCE-----GGVPQRPVVVTRKKLVNAIQKLSK--142
1083 AAGPYVGGLEEFQQLVLDVNNMIQKLYLPQGLLEQ-VSIPDDDLVQATFQQLQKPP 1141
143 -----LKSGPGWVTHGMAGCKSVLAABAVRDHSLLEGCFPGGV---182
1142 SPARPRLLQDTVQXLMPLPHGRSLVLT--GSGGQKTAFLASLV---SALQA--PDGAKVA 1194
183 -----HWVSVGKQKSGILMKLONLCTRLDQESFQRLPLNIEBAKDRILMLRKHPR 237
1195 XLVPFHP-SGARPQGLALTLLRLCTYLRGOLKEPGALPSTYRSLVWELQORLLPKSAE 1253


```
Qy 418 FVNSLLFCDRNG-----KSFYRLHDLQVDFELTEKNSQLODLHKKIITQFQ 465
Db 1463 -----ASHSGNPFPLCPAYLVQSLRLLGEGVPERGARCLSLDGPRTTIKRYG 1514
Qy 466 R-----YHQPHLTSPQEDCMYWNF-----LAYHMASAKHKLCA 502
Db 1515 KRGLGKTAHVLIAAHLWKTCDDPASG-----TFRSCPEALKDPLPHLLQSGNHGLLAE 1569
Qy 503 LMFSLDWIKAKTELVG-----PAHLIH--EFVEYRHILDEKDCAVSNFOEPLSLNGHL 554
Db 1570 FLTNLHVAAAYLE-VGLVPDLLEAHVLYASSKEANOKLPAADVAV---PHTFLRQOASL 1625
Qy 555 LGRQPPNIVOLGLCEPETSVCQAKLOAKQVDNGMYLEWINK-----KNITNLSRLV 610
Db 1626 LTQYPL-LLQQAASQSPESVPCQAPLLTQRWHD--QFTLKNWINKPQLKGQOOLS-LT 1681
Qy 611 VRPHTDVYHACFSDGRIASCADYTLQVFAETGEX----- 649
Db 1682 MSSSPTAV--AFSPNGORAAVGTASGTIYLLNLKWTQEEKAVSGCDGSSFAFLSDTA 1738
Qy 650 -----LLEIKAEDEVLCAPSTDDRIATCSVDKVKIWNMTG 689
Db 1739 LFTTDPGHLEWLDLQHCWVFTKAHQYQITCCCLSPDRRLATVCLGGYLKLMDTVRG 1798
Qy 690 ELVHTYDEHSEQVNCCHFTNSHLLLATGSSDCFLLKMDLNQKECRNTPMFGHTNSVNH 749
Db 1799 QLAFOY-THPKSLNCVAFHPGQ--VVATGSWAGSITFPQADGLKVTYKELGAPGSPVCSL 1855
Qy 750 RFPDDKLLASCADGTLKLDATSNANERKSINVQO-----PFLNLED-----POEDME 798
Db 1856 AFNPKGKIVAVGRIDGTVELW-AMQEGARLAAPPAQCGCSAVLFLHAGORFLTAGED-- 1912
Qy 799 VIVKCCSWGA-----DCARIMVA-AKNKIFLFDHTSGLLG 833
Db 1913 --GKAQWSGFLGRPGCLGSLPLSPALSVLNPDPDQVAVGYREDGINITYKI--SSGSQ 1969
Qy 834 ELHTGHSTIOYCDFSPONHLAVVALSQYCVELWNTDSKSVADCRGHLNVHGWMSFSD 893
Db 1970 PQH-----QELNAVSAI-----VLSPSVLYSGAEDGSL---HGWMEKGD 2007
Qy 894 GSSFL-----TSDDOTIRLWETKKVCKNSAVMLKQEVVDVVFQE 932
Db 2008 SLHSLMLLSRYQKPVGLAASRELMANASDFVRLMFRQ----- 2047
Qy 933 NEWVLAVDHIHRLQJ-----INGRTQIDYLTQAVSCCLSPHLQYIAG--DENGAI 985
Db 2048 ----LLTQPHVHVELPCCAEIRH-----EGPVCCSPSPDGGILATAGRDRLLC 2095
Qy 986 EILELVNNRIFOSRQ--HKVTWHIOFTADEKTLISSDDABIQVN-----WQDKCIP 1039
Db 2096 WDMKTAQAPLLIHTFSSCHRDWITGCAWTKD-NILVSCSSDGSVGLMWPEAQOOLQO--- 2151
Qy 1040 LRGHQETVKDFRLKNSRLSSWSPDGTVKVYNNIITGNKEKDFVCHQGTVLSCDISHDATK 1099
Db 2152 FSGHQSAVSAV-VAVEEHIVSVSRDGTLLKWD-----HOGVELTSIPAHSGPI 2198
Qy 1100 FSSTSA-----DKTAKIWSFDLLPLHLRGNHGNVCRCSAFSDVS 1139
Db 2199 SQCAAALEPRPGQPGSSELLVTVGLDGTATKLWHPLLVCQIRTLQHGSGPVTAASAEAS 2258
Qy 1140 TLATGDDNGEIRIWNVSGELLHLCAPLSEGAATHGGWVTDLCPSPDGMILISA--GG 1197
Db 2259 GLLLTSDS-SVLQWQIFK-EADSYKPRSSVA-----ITAVAWPDGSMVVSNGEAG 2309
Qy 1198 YIKWNV-----VVTGE-----SSOTFYNTGN 1219
Db 2310 ELTLWQQAQAVATAQAQPVRSVHLIWTYSANSFFVLSAN 2346
```

RESULT 13

US-09-184-445-4

; Sequence 4, Application US/09184445

; Patent No. 6174703

```
;
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 6174703el Genes Encoding Telomerase Protein
; NUMBER OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/184,445
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/751,189
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2629 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-184-445-4
```

Query Match 6.08; Score 399.5; DB 3; Length 2629;

Best Local Similarity 21.48; Pred. No. 2.5e-26;

Matches 290; Conservative 179; Mismatches 470; Indels 419; Gaps 62;

```
Qy 120 PQRVVFVTRKLVNALQKLSKLGEPGWTHCMACGK-----SVLAARAVRDHSL 174
Db 1152 PRRP-----RLQDTVQQLLPHGRISLV--QACQGTAFSLASVAKVPDQ--- 1199
Qy 175 EGCFFPGVHWV-----SVGQDKSGLLMKLNCTFLDQDESFSORLPINTEAKDRLIL 230
Db 1200 ----PNEPFFVFFHFAARPDQCLALNLRRLCTHLRQKLGELGALSFTYRGLVWELQK 1255
Qy 231 MLRKPRSL-----LILD-----DVWDSWLKAFDSQCQILLTRDKSVTDSV 273
Db 1256 LLLKFAQSLQPAQTLVLIIDGADKLVDNRNGQLISDWIPKSLPRVHLVLSV---SDSG 1311
Qy 274 MGP-----KYVPVVESSLGKEG---LEILSL-----FVNMKKADLPQAHSTI 314
Db 1312 LGETLQSQGAVVVALGSLVPSSRAQLVREELALYKRLSESPFNQMRLLAKQSSIL- 1370
Qy 315 KECKGSLVWSLIGALIRDP-----PWRWEYVKOLQNKQPKRKSS 357
Db 1371 ----PVLHLVTDYLRFLTYEQVSRRLRTPATLPLLIQHILSTLEQ----- 1415
Qy 358 SYDVEALDEAMSISVEMLRDIDKVVYTDLSILQKDVKVTKVLCLWDMETEVEEDILQE 417
Db 1416 -HGHVLPQALT-ALEVTR-----SGLTVDQLHAULT---WLILPKETKSWEVLA- 1462
Qy 418 FVNSLLFCDRNG-----KSFYRLHDLQVDFELTEKNSQLODLHKKIITQFQ 465
Db 1463 -----ASHSGNPFPLCPAYLVQSLRLLGEGVPERGARCLSLDGPRTTIKRYG 1514
Qy 466 R-----YHQPHLTSPQEDCMYWNF-----LAYHMASAKHKLCA 502
Db 1515 KRGLGKTAHVLIAAHLWKTCDDPASG-----TFRSCPEALKDPLPHLLQSGNHGLLAE 1569
```


